A projector-camera setup for geometry-invariant frequency demultiplexing

Citation: Vaquero, D.A. et al. "A Projector-camera Setup for Geometry-invariant Frequency Demultiplexing." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2009, IEEE, 20-25 June 2009. 2082–2089. Web. ©2009 IEEE.

As Published: http://dx.doi.org/10.1109/CVPR.2009.5206614

Publisher: Institute of Electrical and Electronics Engineers

Persistent URL: http://hdl.handle.net/1721.1/73468

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

A Projector-Camera Setup for Geometry-Invariant Frequency Demultiplexing

Daniel A. Vaquero*
UC Santa Barbara
daniel@cs.ucsb.edu

Ramesh Raskar†
MIT Media Lab
web.media.mit.edu/~raskar

Rogerio S. Feris
IBM Research
rsferis@us.ibm.com

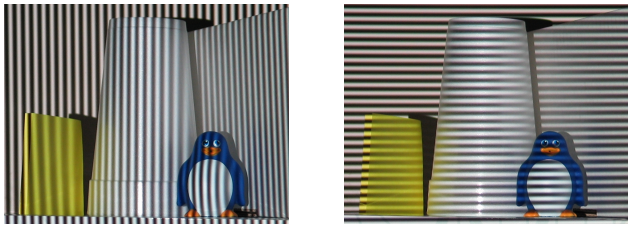
Matthew Turk
UC Santa Barbara
mturk@cs.ucsb.edu

Abstract

Consider a projector-camera setup where a sinusoidal pattern is projected onto the scene, and an image of the objects imprinted with the pattern is captured by the camera. In this configuration, the local frequency of the sinusoidal pattern as seen by the camera is a function of both the frequency of the projected sinusoid and the local geometry of objects in the scene. We observe that, by strategically placing the projector and the camera in canonical configuration and projecting sinusoidal patterns aligned with the epipolar lines, the frequency of the sinusoids seen in the image becomes invariant to the local object geometry. This property allows us to design systems composed of a camera and multiple projectors, which can be used to capture a single image of a scene illuminated by all projectors at the same time, and then demultiplex the frequencies generated by each individual projector separately. We show how imaging systems like those can be used to segment, from a single image, the shadows cast by each individual projector – an application that we call coded shadow photography. *The method is useful to extend the applicability of techniques that rely on the analysis of shadows cast by multiple light sources placed at different positions, as the individual shadows captured at distinct instants of time now can be obtained from a single shot, enabling the processing of dynamic scenes.*

1. Introduction

In computer vision, many active illumination techniques employ projector-camera systems to facilitate the extraction of useful information from scenes. These approaches usually rely on the careful choice of an illumination pattern to be projected onto the objects. The captured image is a function of the projected pattern and its interaction with the objects in the scene; as the projected pattern is known, it is possible to exploit this information to recover properties of the imaged scene. Figure 1(a) illustrates this point. It shows objects imaged under the illumination of a projector placed at the left hand side of the camera. The projector sends vertical white stripes to the scene, but the observed patterns in the image vary according to the local orientation of the objects. Their local frequencies are useful cues for recovering surface orientation.



(a) (b)

Figure 1. The frequency of observed patterns is sensitive to the geometry of the scene. (a) Vertical stripes projected from a projector placed at the left hand side of the camera. Notice the different frequencies on the two slanted planes (left and right), and curved lines on the penguin; (b) By projecting horizontal stripes, the frequency of the observed patterns is geometry-invariant.

In this work, we follow an opposite direction. Rather than exploiting variations in the projected patterns due to depth and orientation changes, we show how a projector-camera setup can be built in a way that the imaged patterns are the same across the scene, no matter what the geometry of the objects is (see 1(b)). Our method is based on a strategic alignment of the projector-camera configuration, and on a particular choice of the projected patterns. The technique is derived from a key observation from the epipolar geometry of two cameras, and the fact that a projector can be understood as a dual of a camera, sharing the same geometric characteristics [12].

We then demonstrate the usefulness of this property for demultiplexing frequencies of patterns simultaneously projected by multiple projectors. In this process, the goal is to, given a single image of a scene illuminated by multiple projectors, be able to determine the patterns and frequencies observed at a given region of the image. Consequently, the demultiplexation of patterns allows us to infer which projec-

*Part of this work was conducted at Mitsubishi Electric Research Labs (MERL) and IBM Research.

†Part of this work was conducted at Mitsubishi Electric Research Labs (MERL).

978-1-4244-3991-1/09/$25.00 ©2009 IEEE

tors are casting shadows along a given region. As the observed frequency is invariant to the geometry of the objects in the scene, the complexity of the demultiplexing process is greatly decreased, since the frequencies to be searched for are fixed and known.

As an application for the frequency demultiplexing scheme, we propose a technique called *coded shadow photography*. It determines a segmentation of the imaged scene, where the image is partitioned according to the shadows cast by the multiple projectors, allowing us to query, for a given region, which projectors are casting shadows over it. We describe a proof-of-concept implementation of a coded shadow photography setup, and present experiments that show how the choice of system parameters affects the results.

Finally, we point out that this method can be employed to extend the applicability of techniques that rely on the analysis of shadows cast by multiple light sources placed at different positions. Those methods usually take multiple pictures of the scene at different instants of time, with only one light source being triggered during the capture of each image. This often brings limitations for scenes with moving objects. Coded shadow photography can be employed to obtain the shadows cast from multiple projectors with a single shot, enabling the processing of dynamic scenes. As an example, we present an experiment where we apply the multiflash algorithm for finding occlusion boundaries [13] having as input the shadows obtained using our method.

1.1. Contributions

The main contributions of this work are:

- We propose a projector-camera setup for which the imaged frequency of a strategically chosen projection pattern is invariant to the orientation and depth of objects in the scene;
- A frequency demultiplexing scheme based on the aforementioned setup;
- An application of the frequency demultiplexing scheme for segmenting shadow regions from a single image captured under the illumination of multiple projectors. We call this *coded shadow photography*.

1.2. Organization

This paper is organized as follows: we begin by reviewing related work in Section 2. In Section 3, we show how a projector-camera setup for which the imaged frequency of projected patterns is invariant to the scene geometry can be achieved. In Section 4, a frequency demultiplexing scheme based on a multiprojector setup is presented. In Section 5, we introduce the *coded shadow photography* technique, which employs our setup to demultiplex, from a single image, the shadows cast by multiple projectors at the same time. We also present experiments that demonstrate how various system parameters impact the results. In Section 6, we describe an application that could benefit from a coded shadow photography scheme: occlusion boundary detection using multiflash imaging. Finally, in Section 7 we discuss possibilities and limitations of the method, and in 8 we present our conclusions and directions for further research.

2. Related Work

Many computational photography methods have exploited the variation of capture parameters (such as exposure [3], focus [20], aperture [5], and viewpoint [15]) to extract information from scenes and/or produce a new enhanced photograph. More recently, great attention has been devoted to techniques that process images by *coding* these parameters. Examples include coded exposure for motion deblurring [11], coded aperture for passive depth estimation [8], and coded viewpoint for active 3D reconstruction [17].

More related to our approach, coded structured light techniques have been studied for a long time in computer vision. Temporal coding methods [10, 6] exploit time-varying patterns of projected light to recover depth with high quality, but are not suited for dynamic scenes. Spatial coding methods [18] handle object motion by utilizing a coded light pattern that varies spatially. However, innacurate results are produced at depth discontinuities, as local surface smoothness is assumed. Spatial-temporal coding [2, 19] combines the advantages of both approaches, assuming object motion coherence. Viewpoint coding [17] relies on multiple cameras for coding structured light, allowing excellent depth reconstruction results without making any spatial or temporal continuity assumptions about the scene. Our approach is related to these methods in the sense that we also project coded structured light onto the scene. However, we deal with a different problem – how to keep the frequency of the pattern invariant across object orientation and depth variations, facilitating frequency demultiplexing when multiple projectors are used at the same time.

Coded illumination through multiflash imaging [13] was proposed for the detection of occlusion boundaries in complex scenes. The approach is based on strategically positioning flashes to cast shadows along depth discontinuities in the scene. By detecting the shadows, occlusion boundaries can be reliably marked. This approach is not capable of handling dynamic scenes well, as the illuminants are triggered at different instants of time. In contrast, our frequency demultiplexing technique allows us to code the shadows in the spatial domain, thus being suitable for occlusion boundary detection in motion.

Most frequency analysis approaches in computer vision

are related to shape from texture methods [14]. These techniques take advantage of the texture/frequency variation in the scene due to perspective distortions in order to extract the shape of surfaces. Active techniques [7] that project patterns with fixed frequencies and analyze their variation across the image have also been proposed for depth recovery. Our approach takes the opposite path: rather than exploiting the projected frequency variation due to depth and orientation changes, we show how the frequency can be made invariant across the scene – a property that can be very useful in other applications, as discussed in this paper.

3. Geometry-Invariant Frequency Projection

Consider a setup composed by a perspective camera and a perspective projector, which projects a sinusoidal pattern $f(x, y) = \frac{h}{2}[1 + \cos(\omega_1 x + \omega_2 y)]$ onto the scene, where ω_1 and ω_2 denote the angular frequencies of the sinusoid, in radians per pixel, and h is the amplitude of the mask (in our experiments, h was set to 255, as we used 8-bit images to represent the masks). As motivated in the introduction, surfaces in the scene can have many possible orientations (for example, tilted planes and curved objects) and can be at different depths. In general, due to these variations the sinusoidal patterns observed in the image captured by the camera may differ from the projected pattern in frequency and phase, as in Figure 1(a). As we will discuss in Section 4, this variability can be a nuisance in some applications.

We exploit a particular epipolar geometry case to overcome this issue. Consider two perspective cameras pointing at the same direction, having parallel optical axes orthogonal to the baseline, which is aligned with the horizontal coordinate axis. This configuration is known in the stereo vision literature as the *canonical configuration* [4]. For this arrangement, the epipolar lines are parallel, and a row in one of the image planes corresponds to the same row in the other image plane. Figure 2(a) illustrates this. If we replace one of the cameras by a perspective projector, the same result holds, since a projector has the same geometry [12]. Thus, if we project horizontal patterns from a projector aligned with the camera along the horizontal coordinate axis, each projected horizontal line will be imaged at the same horizontal line in the camera. Therefore, the frequency of the observed sinusoidal patterns will be insensitive to variations in the shape of the objects. In practice, if the camera and the projector have different resolutions or different focal lengths, the size in pixels of the imaged lines can differ from the size of the projected lines, but the frequency of the observed sinusoidal pattern is still independent of the geometry of the objects in the scene.

The above reasoning suggests that a projector-camera setup for which the frequency of the observed patterns is geometry-invariant can be built by strategically choosing two elements:

- **Projector-camera alignment**. The projector should be placed in the same plane as the camera, such that they point at the same direction and their optical axes are parallel;

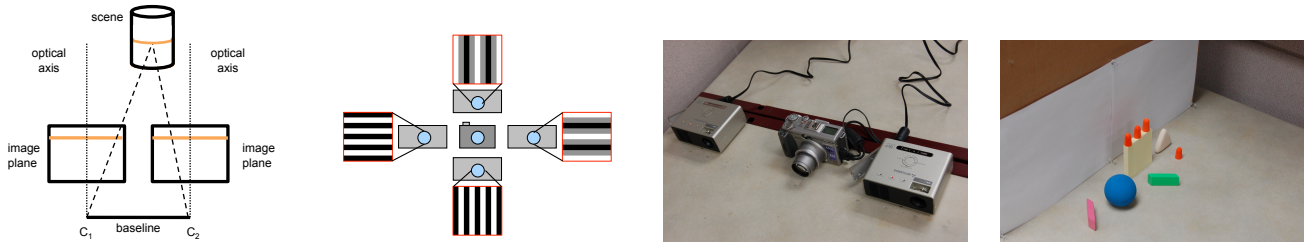
- **Projected patterns**. Patterns parallel to the direction of alignment between the camera and the projector should be projected. For example, for a projector placed to the left or to the right of the camera, it is best to project sinusoids with $\omega_1 = 0$ (horizontal stripes); for a projector placed above or below the camera, it is preferable to use vertical patterns ($\omega_2 = 0$).

By designing a setup with those properties, the frequency of the observed patterns will be invariant to the geometry of objects in the scene, as in Figure 1(b).

4. Frequency Demultiplexing

Let us now consider the problem of frequency demultiplexing in a multi-projector, single-camera setup. Suppose that all projectors send a distinct sinusoidal pattern onto the scene at the same time, and the goal is to analyze the frequencies of observed patterns in the image taken by the camera in order to determine the regions being illuminated by each individual projector. In general, variations in the observed frequency due to the geometry of objects in the scene make it very difficult to detect which frequency came from which projector. On the other hand, if we use the geometry-invariant setup described in the previous section, then the complexity of the problem is greatly reduced, as the observed frequencies are fixed and known. A basic setup for this purpose would consist of a camera and multiple projectors placed in the same plane, such that each projector-camera pair satisfies the conditions described in the previous section. Also, the frequencies of the sinusoids should be distinct among the projectors that share the same direction of camera-projector alignment. Figure 2(b) illustrates an example of a setup that meets these requirements, by projecting sinusoids with frequencies of π and $\pi/2$ radians per pixel.

Given the single image captured using multiple projectors, and the frequencies of the observed patterns, the objective is to, for each projector p_i, determine the image regions that contain sinusoids with the frequency imprinted by p_i. Gabor filters [1] can be used to detect regions where a texture with a specific frequency is present. If $p_1, p_2, \ldots, p_n$ are the n projectors, let ω_{obs_i} be the observed frequency of the sinusoid from projector i. Create n Gabor filters G_i tuned to detect frequencies ω_{obs_i}, $i \in \{1, \ldots, n\}$. Let S be the grayscale image taken using a single shot. For $i \in \{1, \ldots, n\}$, compute H_i, the result of applying the filter G_i to the image S. The filtered images H_i should have low values in regions where the frequency ω_{obs_i} is not present.



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

Figure 2. Geometry of the capture setup. (a) correspondence between epipolar lines in the canonical stereo configuration, considering perspective cameras; (b) sample configuration with a camera and four projectors; (c) experimental table-top prototype; (d) the scene being imaged.

We observe that the problem of segmenting the regions depending on the present frequencies can be understood as a general texture segmentation problem, where the textures being searched for are known a priori (patterns that can be decomposed into a sum of sinusoids, where one of the sinusoids has frequency ω_{obs_i}). We have described a detection algorithm using Gabor filters, but more sophisticated segmentation methods suitable for this task could be applied.

5. Coded Shadow Photography

The solution to the frequency demultiplexing problem from the previous section can be interpreted as a shadow segmentation procedure, where the goal is to determine the regions that are in shadow for each individual projector, starting from a single image captured while all projectors simultaneously cast shadows. The frequency of each sinusoid is a code that allows us to identify regions illuminated by each projector, and the geometry invariance property facilitates the shadow demultiplexing process. We call this application *coded shadow photography*.

In order to demonstrate the ideas, we present experiments carried out using a proof-of-concept setup. We have built a table-top prototype of the capture setup using a four-megapixel Canon G3 digital camera and two Mitsubishi PK20 pocket projectors (native resolution of 800x600 pixels). The projectors have been placed to the left and to the right of the camera, with a camera-projector baseline distance of roughly 17cm. Figure 2(c) shows a picture of our setup, and Figure 2(d) shows the scene being imaged. The objects are at a distance of about 1.2m from the camera, some of their surfaces are curved, oblique occlusion boundaries are present, and a few objects occlude other objects. In order to demonstrate the geometry-invariance property while avoiding frequency variations due to scene albedo, a scene for which the albedo is locally constant in most places was chosen (see Section 7 for a discussion on the effects of texture and possible alternatives to address them).

5.1. Proof-of-concept

To illustrate the entire shadow demultiplexing process, consider the image in Figure 3(a), captured using our setup. It shows a color image of a green eraser in front of a yellow background. Figure 3(b) shows the same image in grayscale. The left projector projects a sinusoid with frequency ($\omega_1 = 0$, $\omega_2 = \pi$) rad/pixel, while the frequency for the right projector is ($\omega_1 = 0$, $\omega_2 = 2\pi/3$) rad/pixel. Notice that only one of the patterns is present in the shadowed regions, while both patterns appear in other regions. As mentioned in Section 3, the frequencies of the observed sinusoids may differ from the frequencies of the projected sinusoids, due to differences in resolution and focal length of the projectors and the camera. For a given projector, if a frequency $\omega_{projected}$ is projected, a frequency $\omega_{observed} = k \cdot \omega_{projected}$ is observed. The observed frequency can be determined by taking a picture of a blank sheet of paper under the illumination of that projector only, locating a cycle of the sinusoidal pattern and then counting the number of pixels spanned by the cycle. The new frequency $\omega_{observed}$ will be equal to $\frac{2\pi}{\#pixels}$. An alternative to counting the number of pixels is to have a predesigned filter bank tuned to detect a few sample frequencies, run the image through the filters in the filter bank, and choose the frequency that corresponds to the filter of maximum response. For the imaging setup used to capture the image in Figure 3(a), the observed frequencies that correspond to the projected frequencies of π and $2\pi/3$ are approximately $4\pi/7$ and $8\pi/21$, that is, the multiplicative factor k is roughly $4/7$.

Figures 3(c-d) show the output of the application of two Gabor filters to the image in 3(b), tuned to detect the frequencies $4\pi/7$ and $8\pi/21$. The intensities are normalized to be in the $[0, 1]$ interval, where darker regions indicate lower response. In order to segment the shadow regions, we applied a simple thresholding operator to the Gabor filter outputs, by selecting the pixels with response values lower than 0.15. The segmented regions are shown in Figures 3(e-f). However, the regions are noisy, even for other thresh-

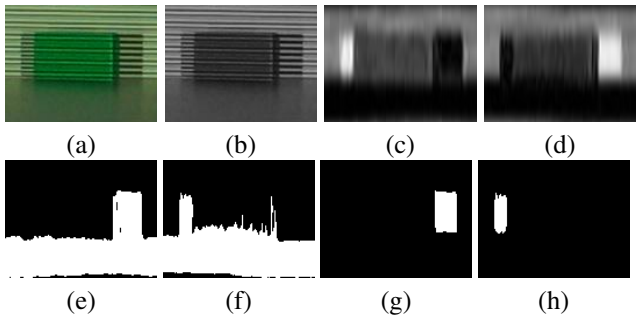


Figure 3. Shadow decoding. (a) color image of an eraser, captured using two projectors, one to the left and one to the right; (b) grayscale version of (a); (c) normalized output of the Gabor filter tuned to detect the frequency of the left projector (darker values indicate lower response); (d) normalized output of the Gabor filter tuned to detect the frequency of the right projector; (e) regions of low response in (c); (f) regions of low response in (d); (g) segmented shadows for the left projector, by considering only the pixels in (e) that have a large difference between their intensities in (c) and (d); (h) same as (g), for the right projector.

old values. Another interesting observation is that, in the shadow regions for one of the projectors, the filter response for the other projector is usually much higher. By filtering the results from Figures 3(e-f) by selecting only the pixels with a large difference between the responses of both Gabor filters, the detection was much more robust, as shown in Figures 3(g-h). In the result, if a given pixel had responses R_1 and R_2 in Figures 3(c-d), it was considered as a pixel with large difference between the responses of the Gabor filters if $|R_1 - R_2| > 0.45$.

5.2. Effects of Capture Parameters

In this section, we evaluate how changes in some of the capture parameters impact the final results. Variations in the camera-projector baseline distance lead to changes in the size of the shadows, and can cause shadow detachment for thin objects. Figure 4(a) shows an image captured using a 17cm baseline, using frequencies of π rad/pixel on the left projector, and $2\pi/3$ rad/pixel on the right projector. Figure 4(b) displays the segmented shadows. We repeated the experiment by placing the projectors on a shorter baseline (12cm), and on a longer baseline (22cm). Figures 4(c-d) show the results for the 12cm baseline, while Figures 4(e-f) show the results for the 22cm baseline.

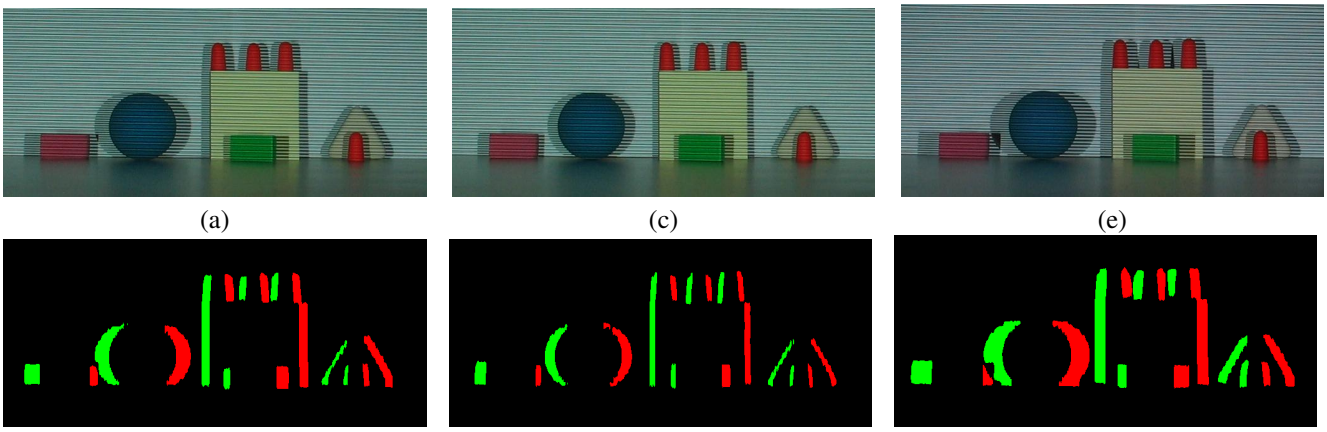
Overall, the segmentation is very good, except for the jaggedness near oblique and curved occlusion boundaries. This is due to the use of Gabor filters to detect the shadow regions. These filters require a spatial support in the vertical direction in order to detect a given frequency, leading to resolution loss near discontinuities in frequency (which appear on oblique and curved edges for horizontal patterns). On the other hand, the segmentation contours are very accurate near vertical occlusion boundaries, as they are not affected by these complications. For the shorter baseline, there is a noticeable increase in failures in the segmentation of the shadows of the triangular object. This is due to the shorter baseline distance, that results in relatively narrow shadows cast by these occluding edges. This reduces the size of the region that contains only the frequency of a single projector, making the detection more difficult for the Gabor filter. Another interesting observation is that with the increase in baseline, the shadows cast over the orange ear plugs begin to detach, but are still captured by our method.

Another important parameter is the frequency of the projected sinusoids. We carried out a set of experiments by using different frequencies on the right projector, while keeping the left projector fixed with a frequency of π rad/pixel. The baseline was 17cm, and we used the frequencies of $\pi/2$, $\pi/3$ and $\pi/4$ rad/pixel on the right projector. Figure 5 displays the results. As the frequency decreases, the resolution of the shadow contours near curved and oblique occlusion boundaries is considerably reduced, due to the spatial support requirement of the Gabor filters. This suggests that it is best to use the highest possible frequencies.

As the experiments demonstrate, the main source of inacuracies in the detected shadows is due to the resolution loss imposed by the Gabor filter near frequency discontinuities. We performed an additional experiment by reducing the resolution of the output of the Gabor filter (by Gaussian filtering and then subsampling) before executing the shadow detection step. This helped to reduce the number of inaccuracies, with the cost of decreasing the resolution of the final output. Figures 6(a-b) show the results of subsampling by factors of 2 and 4, respectively. The images are rescaled to the original size in order to be displayed. Similarly, we performed another experiment by smoothing the output of the Gabor filters using a circular averaging filter (pillbox) of radius 5 before detecting the shadows. The results are shown in Figure 6(c). Notice that the contours are very smooth and free of gaps, at the expense of creating a curvy appearance.

6. Application: Single-Shot Multiflash Photography

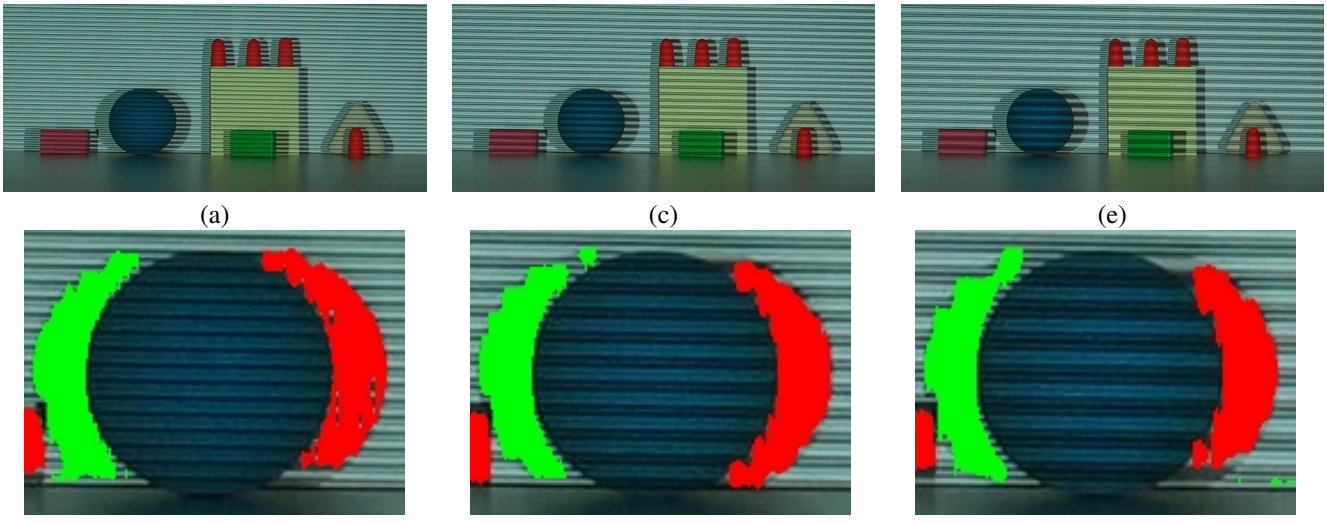
The multiflash technique for occlusion boundary detection [13] uses multiple flashes placed close to the camera, at different locations. A collection of images is acquired, where, for each individual image capture, one of the different flashes is triggered. The algorithm then performs shadow detection in each of the captured images. From the detected shadows and prior knowledge about the relative camera-light placement, the location of the occlusion boundaries is computed. We refer the reader to [13] for more details on multiflash imaging.



(a) (c) (e)

(b) (d) (f)

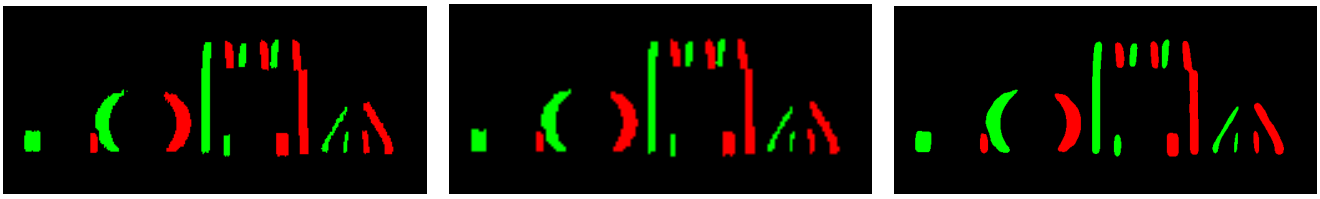
Figure 4. Multibaseline experiment. (a) image captured with a baseline of 17cm; (b) segmented shadows for the image in (a); (c) image captured with a baseline of 12cm; (d) segmented shadows for the image in (c); (e) image captured with a baseline of 22cm; (f) segmented shadows for the image in (e).



(a) (c) (e)

(b) (d) (f)

Figure 5. Experiment with varying frequencies. (a,c,e) images captured with a frequency of $\pi/2$, $\pi/3$ and $\pi/4$ on the right projector, respectively; (b,d,f) zoom to the spherical object with superimposed shadows for (a), (c), and (e). Notice the decrease in accuracy near the sphere's contours as the frequency is reduced.



(a) (b) (c)

Figure 6. Effects of reducing resolution. (a) shadows detected from a lower resolution version (subsampled by a factor of 2) of the Gabor filter outputs; (b) shadows detected from a lower resolution version (subsampled by a factor of 4) of the Gabor filter outputs; (c) result obtained by smoothing the Gabor filter outputs using a circular averaging filter (pillbox) of radius 5 before detecting the shadows.

The method usually produces high quality results for static scenes, but the detection of depth edges in scenes with moving objects is a challenging problem that remains unresolved. Since there is movement during the time interval

between two consecutive captures, the occlusion boundary locations (along with the corresponding shadows) change from image to image. Therefore, it is difficult to reliably find shadow regions, as the images contain misaligned features. Even if a high-speed camera with synchronized flashes is used, it is still not possible to assume that the occlusion boundary locations will coincide in the sequentially captured images.

As shown in the previous section, the coded shadow photography setup can be applied to find shadows from multiple illuminants (projectors) from a single shot, at the expense of resolution. This would enable the use of multiflash photography for moving objects or a moving projectors-camera setup. As it is known that at least three light sources are necessary to cast shadows along all possible orientations of occlusion boundaries in a general scene [16], we built a setup with four projectors, as in Figure 2(b). We captured an image of a scene and applied the shadow demultiplexing algorithm to segment the shadow regions. In the presence of four projectors, we first process the image to filter out either horizontal or vertical frequencies, and then use the same algorithm as before in both resulting images. The occlusion boundary algorithm from [13] was then applied to the detected shadows, in order to find the location of depth edges. Figure 7 shows the results.

7. Discussion

Overall, the experiments show that the frequency demultiplexing scheme can be successfully applied. The geometry invariance property greatly decreases the complexity of the problem, and this reflects in the simplicity of the proposed method. In practice, obtaining a perfect alignment of the camera-projector pairs can be difficult. However, our experiments were performed using approximately aligned projectors arranged manually, and we were still able to obtain good results.

Single-shot coded shadow photography is a promising technique for use in dynamic scenarios. Although the scenes used in the experiments are static, this is merely due to the limited luminance of the pocket projectors we used, requiring an exposure time of 1/6 second during the image capture. Using brighter projectors would enable the use of shorter exposure times. Vision and graphics techniques such as multiflash imaging and relighting could be applied to scenes with moving objects, due to the single-shot characteristic of the method.

In our experiments, we observed that the spatial support requirement of Gabor filters can lead to resolution loss, causing jaggedness at the contours of the segmented regions. Also, small shadows can be missed for the same reason. However, more sophisticated texture segmentation approaches, such as the one proposed in [9], might mitigate this problem. In applications where large resolutions are not essential, subsampling or smoothing the output of the Gabor filters before segmenting shadows can be a viable alternative, as suggested by our experiments. As the sensor and display technologies evolve, higher resolutions are achieved in cameras and projectors. Our method would directly benefit from those improvements, producing higher quality results.

The experiments demonstrate the usefulness of the geometry invariance property, and the technique works well in scenes with locally constant albedo. Although the presence of textured objects can introduce variations in the frequency of the observed patterns, this limitation still can be addressed in some situations. In the presence of high-frequency texture, the frequency of the projected patterns could be decreased at the expense of resolution. Also, while objects with high-frequency texture can cause problems in the visible domain, they may have constant albedo in the infrared domain. This suggests that the use of infrared projectors with an infrared camera might help to address this issue.

As other active illumination methods in computer vision, coded shadow photography is not suitable for use in outdoor environments, when the sun is much brighter than the projected light. Specular reflections can also cause problems, due to the saturation of the captured patterns. The patterns may not be visible due to low albedo of objects in the scene, but the use of brighter projectors would attenuate this issue.

8. Conclusions and Further Research

In this paper, we have proposed a projector-camera setup technique for which the imaged frequency of a strategically chosen projection pattern is invariant to the geometry of the objects in the scene. We have shown the usefulness of this configuration for demultiplexing frequencies projected by multiple projectors at once, and proposed an application called *coded shadow photography*, which is able to recover shadows cast by individual projectors from a single image captured under the illumination of multiple projectors. Finally, we have applied the technique to obtain single-shot occlusion boundary detection using multiflash imaging. As our technique detects shadows from a single image, it may be applied to dynamic scenes.

This work also paves the way for further research topics. We presented an algorithm for single frame processing, but, when video data is available, one could use space-time consistency [19, 2] to combine information from multiple frames, in order to improve the results. Also, the feasibility of projecting other types of patterns might be investigated.

The capture setup is strategically built in a way that the geometry of the surfaces being imaged does not distort the frequency of the projected patterns, in order to facilitate the detection. Following the opposite path, if we choose patterns that exhibit changes on slanted and curved surfaces,

(a) (b) (c)

Figure 7. Using coded shadow photography for single-shot depth edge detection. (a) original image, captured under the illumination of four projectors; (b) shadow detection results; (c) depth edges extracted using the multiflash algorithm.

then the amount of distortion could be exploited to recover the orientation of surfaces in the scene. Indeed, shape from texture methods analyze variations in texture present in objects in order to recover their geometry [14]; the use of projectors would enable the generation of textures over textureless objects, allowing for shape from texture techniques to be applied with the objective of recovering 3D geometry.

Finally, the codification and decodification scheme for shadows proposed in this paper is an inspiration to the investigation of a technique for general multiplexing and demultiplexing of illuminants. In this paper, coded shadow photography generates a binary (lit/non-lit) output for each projector. However, a technique to recover the intensity of the observed patterns (with gray-level / color output) would allow us to decompose a single captured image taken using N coded illuminants into N individual images with lower resolution, as if they had been captured under the illumination of only one of the light sources. We envision the application of this technique in future settings, when improvements in the resolution of cameras and projectors could allow for the use of extremely high frequencies.

Acknowledgements

This work was supported in part by the National Science Foundation under award number IIS-0535293. We used a modified version of the Gabor filter code from `retinal.sourceforge.net` in the experiments.

References

[1] A. Bovik, M. Clark, and W. Geisler. Multichannel texture analysis using localized spatial filters. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12(1):55–73, Jan 1990.

[2] J. Davis, D. Nehab, R. Ramamoorthi, and S. Rusinkiewicz. Spacetime stereo: a unifying framework for depth from triangulation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 27(2), 2005.

[3] P. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH / ACM Trans. on Graphics*, 1997.

[4] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition, 2004.

[5] S. Hasinoff and K. Kutulakos. Confocal stereo. In *European Conference on Computer Vision (ECCV)*, 2006.

[6] E. Horn and N. Kiryati. Toward optimal structured light patterns. *Image and Vision Computing*, 17(2):87–97, 1999.

[7] S. Lee, S. Lee, and J. Choi. Depth measurement using frequency analysis with an active projection. In *International Conf. on Image Processing (ICIP)*, 1999.

[8] A. Levin, R. Fergus, F. Durand, and W. Freeman. Image and depth from a conventional camera with a coded aperture. *SIGGRAPH / ACM Trans. on Graphics*, 2007.

[9] W. Ma and B. Manjunath. Edgeflow: a technique for boundary detection and image segmentation. *IEEE Trans. on Image Processing*, 9:1375–88, Aug 2000.

[10] J. Posdamer and M. Altschuler. Surface measurement by space encoded projected beam systems. *Computer Graphics and Image Processing*, 18(1):1–17, 1982.

[11] R. Raskar, A. Agrawal, and J. Tumblin. Coded exposure photography: Motion deblurring using fluttered shutter. *SIGGRAPH / ACM Trans. on Graphics*, 2006.

[12] R. Raskar and P. Beardsley. A self-correcting projector. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2001.

[13] R. Raskar, K. Tan, R. Feris, J. Yu, and M. Turk. A non-photorealistic camera: depth edge detection and stylized rendering using multi-flash imaging. *SIGGRAPH / ACM Trans. on Graphics*, 2004.

[14] B. Super and A. Bovik. Planar surface orientation from texture spatial frequencies. *Pattern Recognition*, 28:728–743, 1995.

[15] R. Szeliski and H. Shum. Creating full view panoramic image mosaics and environment maps. In *SIGGRAPH / ACM Trans. on Graphics*, 1997.

[16] D. Vaquero, R. Feris, M. Turk, and R. Raskar. Characterizing the shadow space of camera-light pairs. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2008.

[17] M. Young, E. Beeson, J. Davis, S. Rusinkiewicz, and R. Ramamoorthi. Viewpoint-coded structured light. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2007.

[18] L. Zhang, B. Curless, and S. Seitz. Rapid shape acquisition using color structured light and multi-pass dynamic programming. In *Intl. Symposium on 3D Data Processing Visualization and Transmission*, 2002.

[19] L. Zhang, B. Curless, and S. Seitz. Spacetime Stereo: Shape Recovery for Dynamic Scenes. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2003.

[20] L. Zhang and S. Nayar. Projection defocus analysis for scene capture and image display. *SIGGRAPH / ACM Trans. on Graphics*, 2006.