# Generating Spatial Attention Cues via Illusory Motion

Janus Nørtoft Jensen[1*] jnje@dtu.dk
Morten Hannemose[1*] mohan@dtu.dk
Jakob Wilm[2] jaw@mmmi.sdu.dk
Anders Bjorholm Dahl[1] abda@dtu.dk
Jeppe Revall Frisvad[1] jerf@dtu.dk
Serge Belongie[3] sjb344@cornell.edu

[1]Technical University of Denmark [2]University of Southern Denmark [3]Cornell Tech

## Abstract

*For many applications in augmented reality (AR), the user has a much more enjoyable experience if the AR system is able to properly guide the user's attention. In this extended abstract, we explain how to create patterns of light that when projected onto an object are perceived as if the object itself is moving. This can be used as a spatial attention cue. We accomplish this with a calibrated projector-camera setup to synthesize an image from the projector's point of view. This image is filtered to create local phase changes that are then projected back onto the object and perceived as motion. Our method will be shown as a live demonstration at the CV4AR/VR workshop at CVPR 2019.*

(a) Left camera (b) Projector (c) Right camera

Figure 1. Example of a synthesized image from the projector's point of view (b), along with the camera images the intensities are sampled from (a, c).

## 1. Introduction

Visual attention is important as it affects our performance in many visual tasks [2]. To successfully accomplish tasks such as visual search or tele-assistance we need effective spatial attention cues to attract the attention of a user. What constitutes an effective cue is in part determined by the task at hand. A very obvious cue, such as a big bouncing red arrow, might be the best, if the purpose is to alert the user of possible danger. In many other situations, however, more subtle cues might be preferred. This could be in the setting of an escape room game where we would guide players towards the next clue if they are stuck. A very obvious cue could possibly ruin the fun of the game, however a subtle one, which would take longer to notice, could be useful.

As described by Carrasco and Barbot [2], our attention is involuntarily captured by sudden changes in the environment. We seek to exploit this effect by creating apparent motion through light projection. While the human ability to detect motion is not better in the peripheral vision [10], the speed of visual processing does increase in the peripheral vision [3]. When something moves differently compared to the movement of the observer, it becomes a powerful cue to attract attention, especially in the periphery of the visual field [12]. This is referred to as a relative-motion cue. Our idea is to guide attention by creating apparent relative-motion cues in a scene by means of a light projector that modifies the appearance of a physical object to make it look as if it were moving, thereby creating illusory motion.

## 2. Projecting Illusory Motion

Prior work in guiding visual attention has mainly been focused on head-mounted displays and images displayed on a screen. The spatial attention cues used include flickering [16, 17], blurring [8], and color manipulation [9]. Spatial projection has been shown to be more effective than head-mounted displays in providing spatial instructions in assembly tasks [1]. Research in attention cues for spatial projection is however limited. Some use laser projection [14, 7] as a simple cue. Similar to our approach Taki-

*These authors contributed equally to this work.

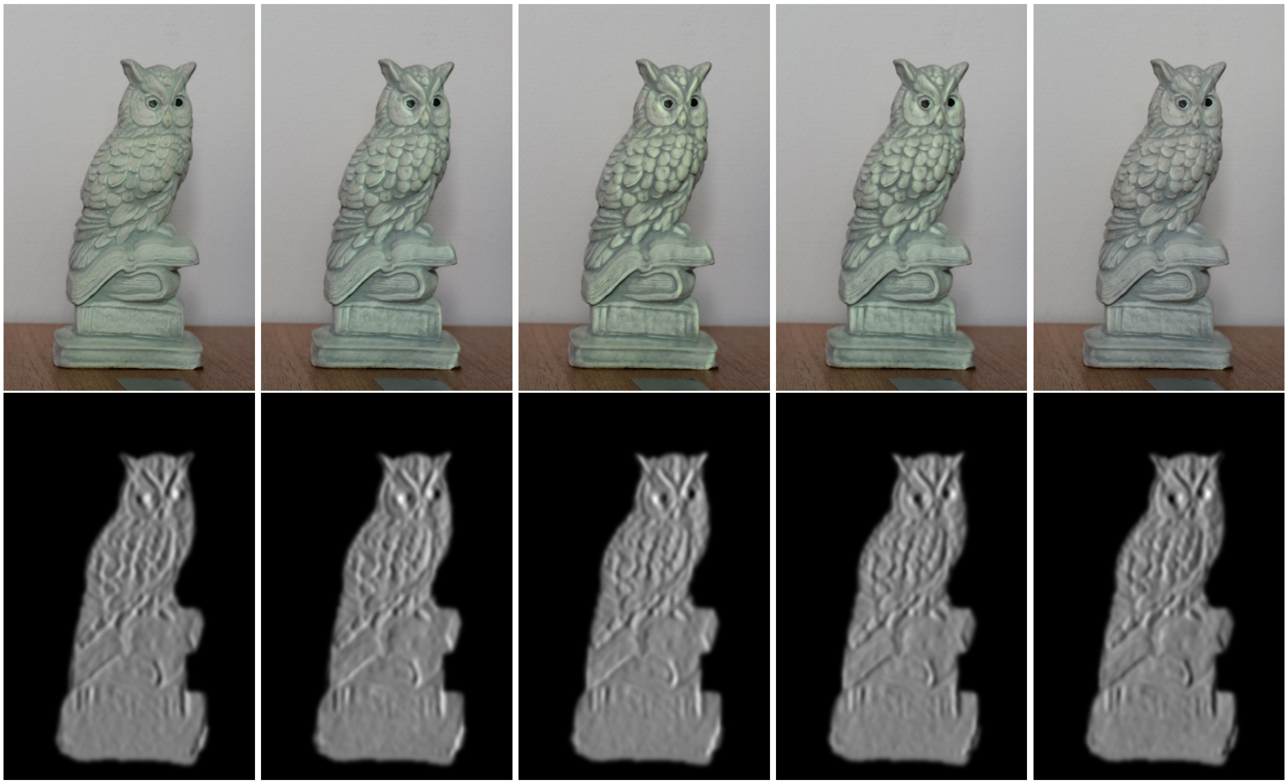

Figure 2. Five frames sampled from the continuous loop of motion our method produces. Top row: Picture of object taken with camera. Bottom row: Image projected by projector to create corresponding picture. Videos showing the effect can be seen at http://people.compute.dtu.dk/jnje/illusory-motion.

moto et al. [15] uses a calibrated projector-camera setup. They modulate color information recorded with the camera and project the result back onto the object. For humans, the sensitivity to color variations declines faster compared with the sensitivity to luminance [6]. It thus seems natural to investigate modifying the luminance instead.

Our approach to create the illusion of motion is based on adapting the work by Freeman et al. [4] to a projector-camera setup. In short, they apply local filters with continuously varying phase over time to the image. This is based on the observation that local phase changes are interpreted as global motion. We synthesize an image from the viewpoint of the projector (Figure 1) and use it as input for their method. The filter response is then projected back onto the object. In Figure 2, examples of the object with the filter response projected onto it are shown along with the images projected.

### 2.1. Projector-camera calibration

We use two cameras and a projector mounted in a fixed setup, and we model the projector and cameras as pinhole cameras with radial distortion. To calibrate the system we encode the projector's pixel coordinates using structured light [13], detect corners in images of a checkerboard and convert these to the projector's pixel space by local homographies [11]. The projector and cameras are then calibrated with Zhang's method [18].

### 2.2. Synthesizing image from projector's point of view

To synthesize an image from the projector's point of view we use structured light to create two 3D-scans of the object - one based on each camera. The resulting point clouds and the pixel intensities associated with the points are then projected back to the projector to form the desired image (Figure 1). Creating two separate 3D-scans enables us to scan all points visible in the projector and at least one camera, thereby getting better coverage of the object. After the points from each scan have been projected to the projector's pixel space they are rounded to the nearest integer pixel. Because the projector has a lower resolution than the cameras, each pixel in the projector contains multiple measurements. We compute medians of these to obtain a single value per pixel.

### 2.3. Creating Illusory Motion

With the image from the projector's point of view, we can directly apply the method of Freeman et al. [4] to this image. As the filter responses contain both positive and negative values, and the projective setting has the constraint of only adding light, we add a constant value to the filter response to make all values positive. The filtered image is multiplied with a mask to restrict light to the object, and the resulting image is projected onto the object. Examples of the projected image and the resulting effect are in Figure 2.

## 3. Future work

We can perhaps utilize the 3D information obtained through our process to make the projected patterns less view dependent. Because we know the 3D position of each pixel in addition to its intensity, we can compute a normal at each point. If we, based on these normals, choose a consistent orthonormal basis at each point, such that the basis changes smoothly over the object [5], and assuming that the object is locally planar, we can project the filter onto this plane, and thereby approximate convolution along the surface of the object instead of in image space.

Our current work has been focused on creating the illusion of motion. Further work is needed to determine under which circumstances it is perceived as motion and to determine its effectiveness as a spatial attention cue. Furthermore, examining how it affects the user's experience of interacting with the AR system, and how our cue compares with using alternative spatial attention cues is important to examine. Reasonable variables to look into would be how much light is necessary to make the effect noticeable and how this varies across the visual field.

## References


[1] Sebastian Büttner, Markus Funk, Oliver Sand, and Carsten Röcker. Using head-mounted displays and in-situ projection for assistive systems: A comparison. In *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, PETRA '16, pages 44:1–44:8. ACM, 2016. 1

[2] Marisa Carrasco and Antoine Barbot. Spatial attention alters visual appearance. *Current opinion in psychology*, 29:56–64, 2019. 1

[3] Marisa Carrasco, Brian McElree, Kristina Denisova, and Anna Marie Giordano. Speed of visual processing increases with eccentricity. *Nature Neuroscience*, 6(7):699, 2003. 1

[4] William T Freeman, Edward H Adelson, and David J Heeger. *Motion without movement*, volume 25. Citeseer, 1991. 2

[5] Jeppe Revall Frisvad. Building an orthonormal basis from a 3d unit vector without normalization. *Journal of Graphics Tools*, 16(3):151–159, 2012. 3

[6] Thorsten Hansen, Lars Pracejus, and Karl R Gegenfurtner. Color perception in the intermediate periphery of the visual field. *Journal of Vision*, 9(4):26–26, 2009. 2

[7] Harald Haraldsson, Doron Tal, Karla Polo-Garcia, and Serge Belongie. Pointar: Augmented reality for tele-assistance. In *CVPR Workshop on Embedded Computer Vision*, Salt Lake City, UT, 2018. 1

[8] Hajime Hata, Hideki Koike, and Yoichi Sato. Visual guidance with unnoticed blur effect. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, AVI '16, pages 28–35. ACM, 2016. 1

[9] Victor A. Mateescu and Ivan V. Bajić. Attention retargeting by color manipulation in images. In *Proceedings of the 1st International Workshop on Perception Inspired Video Processing*, PIVP '14, pages 15–20. ACM, 2014. 1

[10] Suzanne P McKee and Ken Nakayama. The detection of motion in the peripheral visual field. *Vision research*, 24(1):25–32, 1984. 1

[11] Daniel Moreno and Gabriel Taubin. Simple, accurate, and robust projector-camera calibration. In *2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission*, pages 464–471. IEEE, 2012. 2

[12] Dorothe A Poggel, Hans Strasburger, and Manfred MacKeben. Cueing attention by relative motion in the periphery of the visual field. *Perception*, 36(7):955–970, 2007. 1

[13] Carsten Reich, Reinhold Ritter, and Jan Thesing. White light heterodyne principle for 3d-measurement. In *Sensors, Sensor Systems, and Sensor Data Processing*, volume 3100, pages 236–245. International Society for Optics and Photonics, 1997. 2

[14] Björn Schwerdtfeger, Daniel Pustka, Andreas Hofhauser, and Gudrun Klinker. Using laser projectors for augmented reality. In *Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology*, VRST '08, pages 134–137. ACM, 2008. 1

[15] Hironori Takimoto, Katsumi Yamamoto, Akihiro Kanagawa, Mitsuyoshi Kishihara, and Kensuke Okubo. Guiding visual attention based on visual saliency map with projector-camera system. In *HCI International 2017 – Posters' Extended Abstracts*, pages 383–390. Springer International Publishing, 2017. 2

[16] N. Waldin, M. Waldner, and I. Viola. Flicker observer effect: Guiding attention through high frequency flicker in images. *Computer Graphics Forum*, 36(2):467–476, 2017. 1

[17] M. Waldner, M. Le Muzic, M. Bernhard, W. Purgathofer, and I. Viola. Attractive flicker guiding attention in dynamic narrative visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2456–2465, Dec 2014. 1

[18] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, 22, 2000. 2