Invisible gaze direction control in a teleconference

Taeg Sang Cho
Researcher in residence - HP Labs
Massachusetts Institute of Technology

1. Introduction

We are interested in developing a method to direct or control user's gaze in a teleconference environment. This problem is important when the object to focus on, i.e. the person talking, is not obvious from the video stream due to clutters in the scene. In this work, we assume that the region of interest in the video stream has been detected through other means: we focus on signaling the region of interest to viewers with minimal visual artifacts.

One obvious way to draw the user's attention is to place a bounding box around the region of interest. However, the bounding box obstructs the video content, and can be visually disturbing. The computer graphics community developed methods that keep the focus in the region of interest and blurs uninteresting regions [7], or that dynamically change the level of details in stylized renderings according to the level of interest [3, 4]. All these techniques modify the image/video content and *force* the user to look at certain directions, which may not be desirable for viewing experiences.

Recently, Bailey *et al*. [2] introduced an invisible gaze control system leveraging the eye-tracking system. The display flickers the region of interest to grab the user's attention, and as soon as the user directs the gaze to the desired region (detected using the eye-tracker), the flicker stops, thereby not interfering with the viewer experience. We extend the idea of controlling the gaze direction without sacrificing the viewer experience, but also without using an eye-tracker.

The main idea of this work is to make use of the physiological characteristics of human eyes. Photoreceptors that are sensitive to short-wavelength waves, called S-cones, are present only at the periphery of human eyes. If we can design a flicker signal that triggers just the S-cones and not the other photoreceptors, we can, in theory, attract the user's attention to the region of interest, but when the user looks straight at the region of interest, the user shall not perceive the flicker. This visit to HP Labs focused primarily on developing methods to implement this idea.

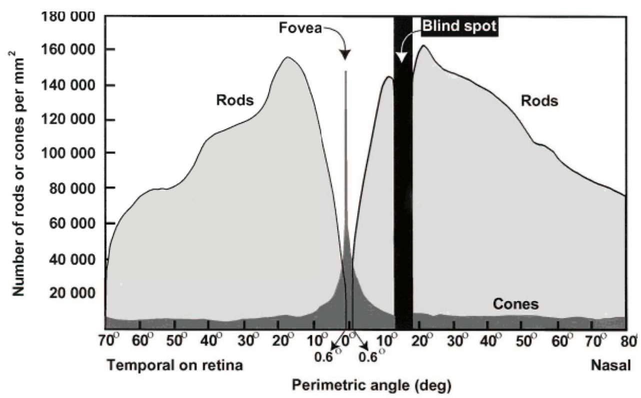

Figure 1. Cones are concentrated at the center of retina (i.e. the foveal region) whereas rods are densely populated in the periphery of retina. (adopted from [5])

2. Background

There are two types of photoreceptors in human eyes: rods and cones. Rods respond to a low dose of illumination (i.e. achromatic), making it suitable for viewing low-light scenes. Cones respond to chromatic light and are less sensitive compared to Rods. Cones are concentrated at the center of retina (i.e. the foveal region) whereas rods are densely populated at the periphery of retina.

There are three types of cones: L-cones, M-cones and S-cones. Different cones are sensitive to different spectrums of light (Figure 2): L-cones are sensitive to long wavelengths (~ 600 nm); M-cones are sensitive to moderate wavelengths (~ 550nm); S-cones are sensitive to short wavelengths (~ 430 nm).

Of particular interest to us are S-cones. S-cones constitute only about $8 - 10\%$ of the cone photoreceptor population in retina [6], rendering the human visual perception less sensitive to blue compared to red or green. Also, S-cones are extremely sparse in the central region of retina, called the S-cone blind spot [8]. The diameter of the S-cone blind spot is about $100\mu m$ or $0.35°$ of retina.

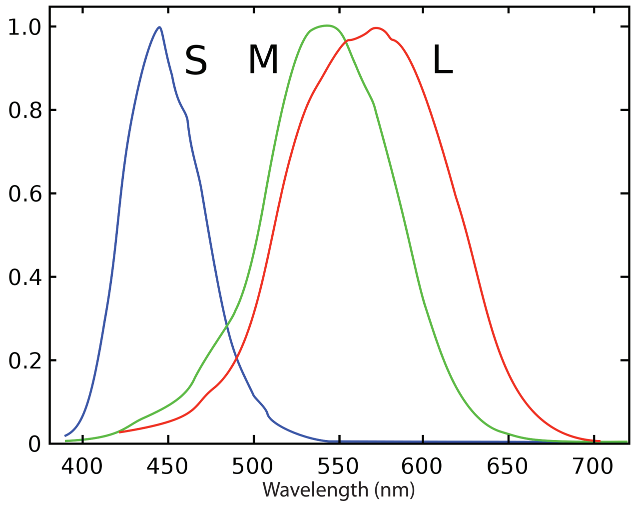

Figure 2. Different cones have different spectral sensitivities.

3. Implementation ideas

Our idea is to make use of the S-cone blind spot. By overlaying the video stream with a small-enough flicker pattern that just stimulates the S-cones and not the other types of cones, we can direct the user's gaze to the desired object/location while being invisible when focused on the desired object.

Since S-cones are sensitive to ~ 430 nm waves, we can overlay the video stream with violet flickers. We tried this idea. Given that the S-cone blind spot is about 0.35° and that a casual user would be about $50cm$ away from the display, the size of the flicker blob should be less than $3mm$. We rendered a video stream with a flickering violet dot, running the flicker at close to 25Hz [1]. However, the flickers were visible when we focused on them. The reason is that the violet on regular displays is *not* the true violet with ~ 430nm wavelength, but only a *metamer* of the violet we need. Therefore, exciting the S-cones using the flickers on regular LCD displays is difficult.

One solution to this problem involves using LED-backlit LCD monitors. Recently, electronics companies, Samsung in particular, released a line of LED TV's that back-lights the LCD using LEDs. Among many benefits that the LED backlight has, an interesting feature we can leverage is that the LEDs are narrow-band. An LED backlight panel consists of red, green and blue LEDs with narrow wavelengths (i.e. the true red, green, and blue instead of their metamers). Figure 3 shows that the blue light emitted from the LED is narrower-band compared to the blue-channel color filter from the LCD. In fact, the spectral content of the blue LED for the backlight panel is quite well aligned with the spectrum to which S-cones are sensitive to Figure 3. We can use

[1] Running the flicker at a higher rate would cause aliasing in videos due to 60Hz display refresh rates

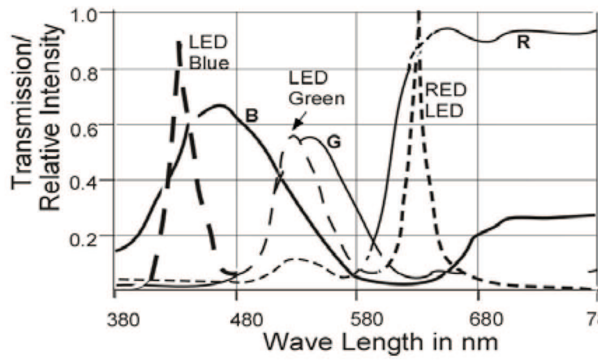

Figure 3. LED spectral characteristics and LCD color filter spectral characteristics (Adopted from [1])

such LED for signaling the region of interest. Since LEDs can be small, we should be able to generate a flicker pattern that is only sensitive to the peripheral S-cones.

References

[1] M. Anandan. Led backlight: Enhancement of picture quality on lcd screen. In *Asian Symposium Information Display*, 2006.

[2] R. Bailey, A. McNamara, N. Sudarsanam, and C. Grimm. Subtle gaze direction. *ACM Transactions on Graphics*, 2009.

[3] F. Cole, D. DeCarlo, A. Finkelstein, K. Kin, K. Morley, and A. Santella. Directing gaze in 3d models with stylized focus. In *Eurographics symposium on rendering*, 2006.

[4] D. DeCarlo and A. Santella. Stylization and abstraction of photographs. In *29th Annual conference on computer graphics and interface techniques (SIGGRAPH)*, 2002.

[5] N. Hadjikhani and R. Tootell. Projection of rods and cones within human visual cortex. *Human brain mapping*, 2000.

[6] R. E. Marc. Chromatic organization of the retina. *Cell Biology of the Eye*, 1982.

[7] G. E. Mitchell. Taking control over depth of field: Using the lens blur filter in Adobe Photoshop CS. Technical report, The light's right studio, 2004.

[8] D. R. Williams, D. I. A. MacLeod, and M. Hayhoe. Punctate sensitivity of the blue-sensitive mechanisms. *Vision research*, 21, 1981.