# Gaze Contingent Foveated Rendering

Sanyam Mehra, Varsha Sankar

{sanyam, svarsha}@stanford.edu

*Abstract*—The aim of this paper is to present experimental results for gaze contingent foveated rendering for 2D displays. We display an image on a conventional digital display and use an eye tracking system to determine the viewers gaze co-ordinates in real time. Using a stack of pre-processed images, we are then able to determine the blur profile, select the corresponding image from the image stack and simulate foveated blurring for the viewer. We present results of a user study and comment on the employed blurring methodologies. Applications for this technique lie primarily in the domain of VR displays where only a small proportion of the pixels that are rendered lie in the foveal region of the viewer. Thus, promising to optimize computational requirements without compromising experience and viewer comfort.

## I. INTRODUCTION

GAZE contingent display techniques attempt to dynamically update the displayed content according to the requirements of the specific application. This paper presents one such technique that exploits the physiological behavior of the human visual system to modify the resolution in the peripheral region while maintaining the resolution of regions in the foveal field of view.

Extension of this technique promises computational savings for rendering on planar and VR displays with expected increase of display field-of-view in the future.

Standard psychophysical models suggest that the discernible angular size increases with eccentricity. Models like [7] predict visual acuity falls off roughly linearly as a function of eccentricity. The falloff is attributed to reduction in receptor density in the retina, as shown in Fig. 1, and reduced processing power in the visual cortex committed to the periphery. [6] suggests only a small proportion of pixels are in the primary field of view, especially for head-mounted displays (HMD). The growing trend towards rendering on devices like HMDs, portable gaming consoles, smartphones and tablets motivates the goal to minimize computation while maintaining perceptual quality.
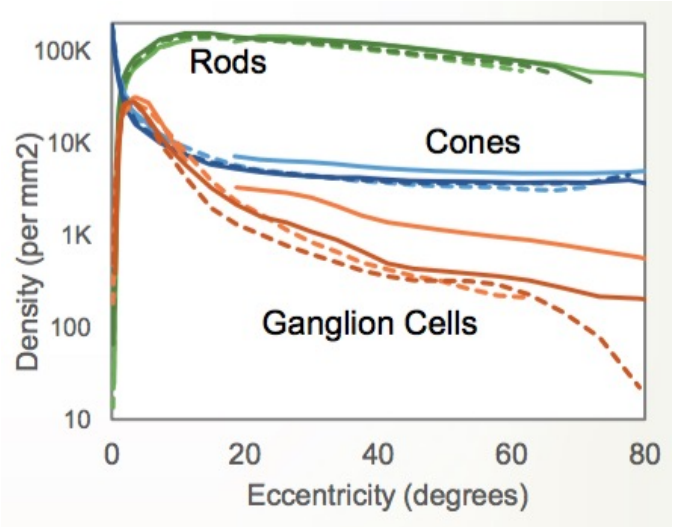


Fig. 1. Receptor density of the retina vs. eccentricity. Adapted from Patney et al. 2016

Given the acuity vs. eccentricity model predictions, the slope characterizing the falloff allows devising blurring methods i.e. angular span and the magnitude of the blur. Section IV presents analysis of the performance of gaze based foveated rendering. The resulting image is expected to appear similar to a full-resolution image, with reduction in the number of pixels required to be rendered, while still maintaining the same perception quality. Section V-E shares results of a user study conducted to evaluate effectiveness of the system with varying parameters, as mentioned above.

Fig. 2 illustrates the practical test setup wherein the gaze location on the screen determines the regions that fall into focus, which in turn dictates the foveated blur. Ideally, the demonstration would require a system with an eye tracking enabled HMD. But, due to lack of readily available hardware, the experimental setup comprises of a platform for a 2D monitor integrated with the EyeTribe eye-tracker and a rendering pipeline that renders pre-processed images.
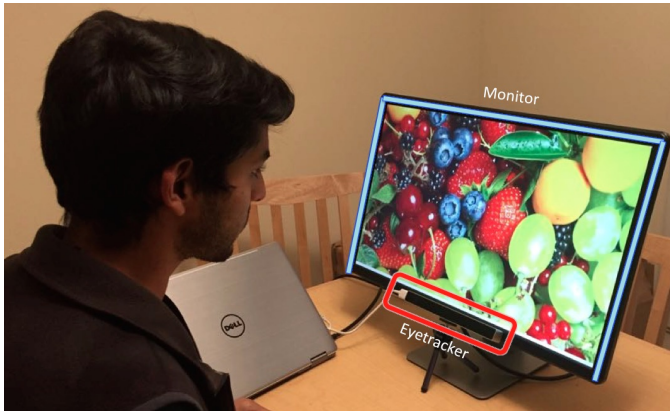
Fig. 2. Setup showing the Eyetribe eyetracker, the 2D monitor and a viewer

## II. Related Work

Some related work in creating foveated rendering algorithms exploits foveation without eye tracking, with the assumption that a viewer primarily looks at the center of the screen [3], or by using a content aware model of visual fixation as in [5], [8]. Such work provides statistical validity across temporal axes and across different users, but fails to account for real-time feedback of viewer's gaze fixation. In [2], they degrade the resolution of peripheral image regions to help in real time transmission of data as well as improve realism of displayed content.

More recent work in this field [4], claims graphics computation optimization by a factor of $5-6$ on a full-HD desktop display. [6] reports the user study they conducted to test foveated rendering for HMDs. Participants judged whether the blurring in a rendered scene was perceptible. Current methods point towards expected computational savings by using the proportion of pixels rendered for the normal vs. foveated blurred rendering cases.

## III. Methodology

The complete process flow of the setup is illustrated in Fig. 3.

### A. Image Stack Pre-processing

The image is loaded and then divided into a grid of sub-images. Due to hardware constraints, this is carried out to prepare a stack of pre-processed images that can be used to simulate real-time foveated rendering. The images are created as per the parameters of the model described in section IV. The grid-dimension is a hyper-parameter that is empirically tuned based on the image/monitor resolution, viewing distance, hardware constraints. A pixel-wise index mask is created and stored, that is later used to select the image to be displayed, based on the sub-image in the grid that the real-time gaze co-ordinates map to.

### B. Gaze Tracking

The EyeTribe eyetracker is linked to the rendering system and it is used to report the viewer's gaze co-ordinates at 30fps. To obtain noise-free reliable gaze-coordinates, each reading of the tracker is processed and classified as a fixation vs. a saccade for the purpose of this experiment. This is done by recording a number of readings over a small time frame (a hyperparameter) and evaluating and classifying the 'momentum' of the eye movement by computing the 'gaze velocity' from consecutive tracker readings. Only readings classified as fixations are passed on to the rendering pipeline. This leads to a trade-off between system latency and final perception quality. Given the experimental setup and hardware specifications, we found using 5 consecutive readings for classification as the optimal option.

### C. Rendering

Once the gaze co-ordinates have been received from the eye tracker, the system matches the value with the index mask, selects the pre-processed image from the image stack and refreshes the image rendered on the screen.

## IV. Foveated Rendering Model

As the acuity falls off with eccentricity, the minimum angular region that can be resolved (Minimum Angle of Resolution - MAR) by the eye increases, as in Fig. 4. Thus, decreasing the resolution of the image with eccentricity according to the increase in MAR will emulate the foveation of the eye. Instead of continuously varying the resolution, it is proposed to discretize in order to maximize computational savings, with the expectation of no resulting perceptual differences. The number of discretized regions involves a trade-off between perceptual quality and computational savings.

The discretization of the MAR function is conservative as it lies below the desired line, and maintains higher frequency than maximum perceptible limit.
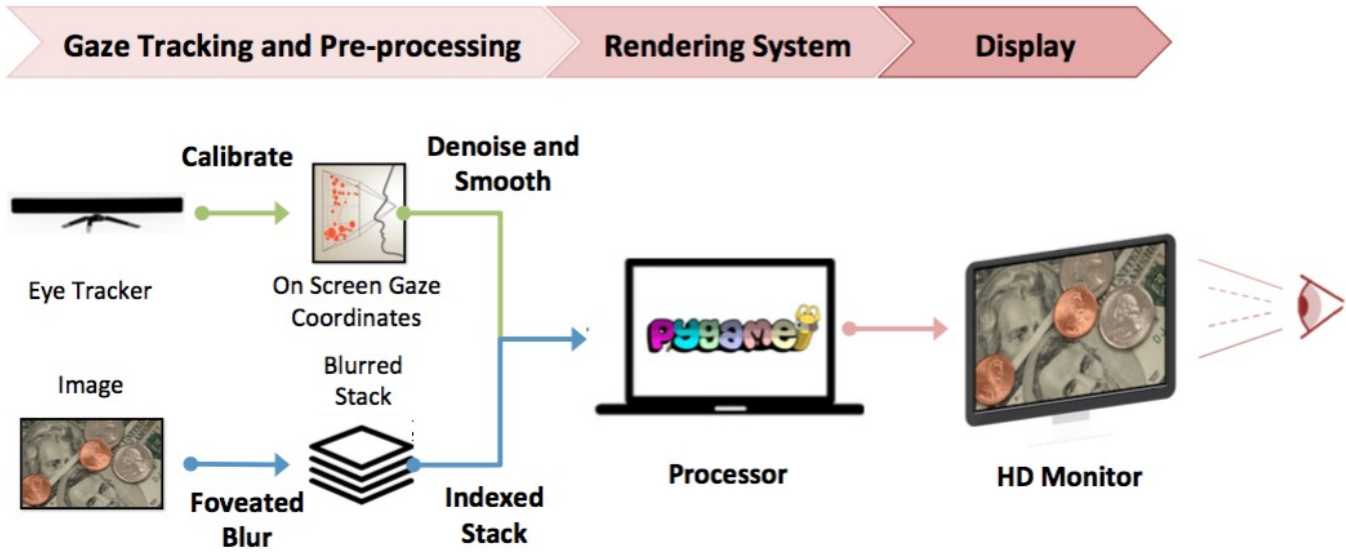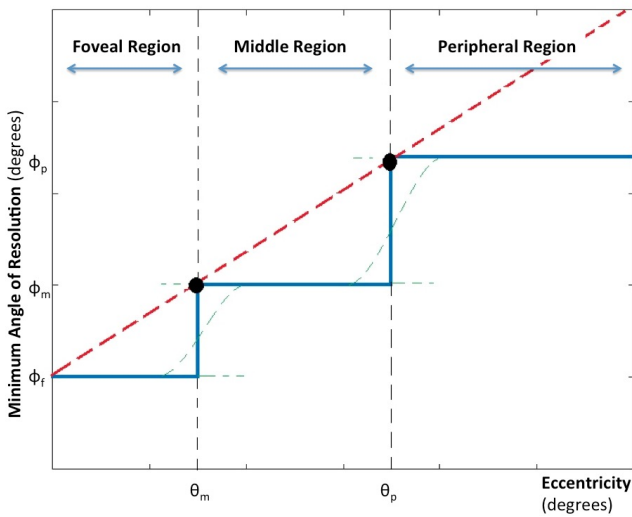
Fig. 3. Implementation flow



Fig. 4. Maximum Angle of Resolution vs. Eccentricity. The red line shows the aspired display behavior. Model can be optimized over number of regions, angular radii $\theta_m$, $\theta_p$, blur magnitudes $\phi_m$, $\phi_p$ The solid blue line and the dotted green line correspond to discrete and progressive blurring respectively

For the purpose of the experiment, the image was divided into three regions; Namely Foveal Region, Middle Region, Peripheral Region. In addition to above-mentioned trade-off, this choice of three regions was guided by hardware restrictions of the rendering system. The beginning of the middle and peripheral regions is marked by the angles $\theta_m$, $\theta_p$. The Peripheral region is rendered with the least resolution followed by the Middle region, while the Foveal region is displayed at maximunm resolution.

Two different approaches were employed to achieve the foveated blur, as discussed below.

*A. Subsampling Foveated Blur*

In this method, the parts of the image in the Middle and Peripheral regions are obtained by subsampling the original image with increasing size of the subsampling filter. Prominent aliasing was observed when displayed and compared to the original image; Certain parts of the image appeared displaced with respect to the original image.

According to [4], The sampling factor ($s$) to be used in each region is determined as follows (for foveal region $s_f = 1$);

$$s_m = \frac{\phi_m}{\phi_f} = \frac{m\theta_m + \phi_f}{\phi_f} \tag{1}$$

$$s_p = \frac{\phi_p}{\phi_m} = \frac{m\theta_p + \phi_f}{\phi_f} \tag{2}$$

where $m$ is the slope of the MAR function. The aliasing was perceivable until small sampling factors were used. But such minimal subsampling offers little advantage in terms of computational savings. Thus, this approach was not included in the User Study.

*B. Gaussian Foveated Blur*

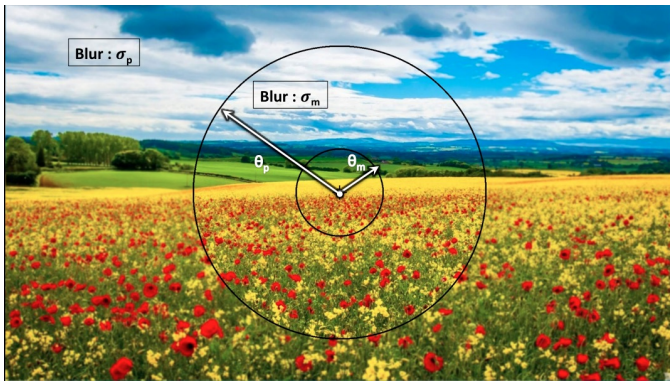This method blurs the middle and peripheral regions using a Gaussian kernel. Three discrete

Fig. 5. Sample image with blurred Middle and Peripheral regions

layers are considered and two different methods were explored; discrete vs. progressive.

The maximum frequency that can be perceived at an eccentricity is directly related to the MAR function. Thus, in the frequency domain, the Gaussian blur to be applied to a region has a standard deviation set to equal the maximum perceivable frequency in that region.

The images used had a resolution of $1920 X 1080$ pixels and the viewing distance was 25 cm. Based on this, the pixel size and pixels per degree ware calculated to be 0.0277 cm and 16 pixels/degree.

As per [1], at $9°$ eccentricity, acuity drops to $20\%$ of the maximum acuity, and at $30°$ acuity drops to $7.4\%$ of the max acuity. Thus, the corresponding maximum perceivable frequency calculated with respect eccentricity and equated to the sigma of the Gaussian blur in the frequency domain. The low-pass filter corresponding to attenuating the frequency components in order to emulate the blur is being estimated as a Gaussian kernel.

Both these techniques were applied by varying the intensity of the blur, discretely and progressively. The resultant images in the two cases were not perceptually discernible. As a result, we present discussion only on the discrete case.

## V. USER STUDY

### A. Participants

The experiment was conducted on ten participants aged between 20 and 35 with normal or 20/20 corrected vision.

### B. Setup

A 24 inch Full HD Monitor with resolution of $1920 X 1080$ pixels (Dell E2414H) was used for display. Sample images with the same resolution were used. The Eye Tribe eye tracker was used to track the gaze of the participants at $30 fps$. The participants were made to view the screen from a distance of 25 cm from the screen.

### C. Experiments

The study was designed to find the maximum foveated blur perceived to be not discernible from the original image, in the absence of abrupt transitions. Two different experiments with the same ten participants were conducted, navigating the space of blur magnitude and the angular radii of the regions.

*1) Experiment 1: Blur Magnitude:* The participants were displayed four different image scenes as separate streams sequentially. In one such sequence, the viewer would start with looking at the original image. The value of standard deviation ($\sigma$) of the Gaussian blur in the middle and the outer regions were varied to produce a stack of blurred images. Each of these images were then displayed alternatively with the original image in order of increasing blur magnitude. The transitions were smooth, with a short exposure to a black scene. The participants were asked to report when they observed any difference.

*2) Experiment 2: Angular Radii:* The participants were displayed four different image scenes as separate streams sequentially. In one such sequence, the viewer would start with looking at the original image. The value of angular radius $\theta_m$ was varied to produce a stack of blurred images. Each of these images were then displayed alternatively with the original image in order of decreasing angular radius. The transitions were smooth with a short exposure to a black scene. The participants were asked to report when they observed any difference.

### D. Sample Images

The four different image scenes that were shown to the users consisted of two natural images, one text image and one binary checkerboard image. The motivation to select this set of images was to understand the variation in response to most widely applicable natural images vs. more structured gaze pattern in text images vs. a high frequency image. Luminance, colour information and frequency profile of content are expected to affect the perception quality. The original sample images used for the experiments are shown in Fig. 6.
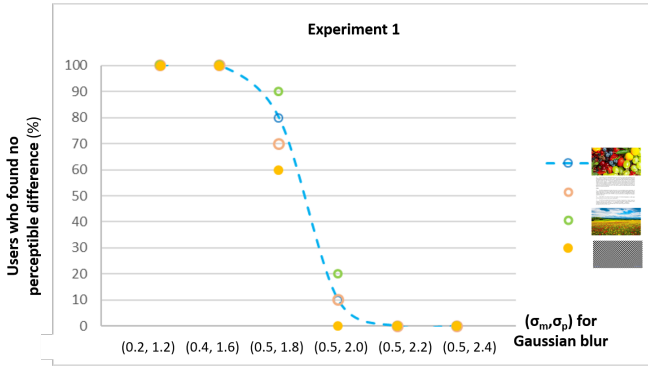
Fig. 6.    Sample images.



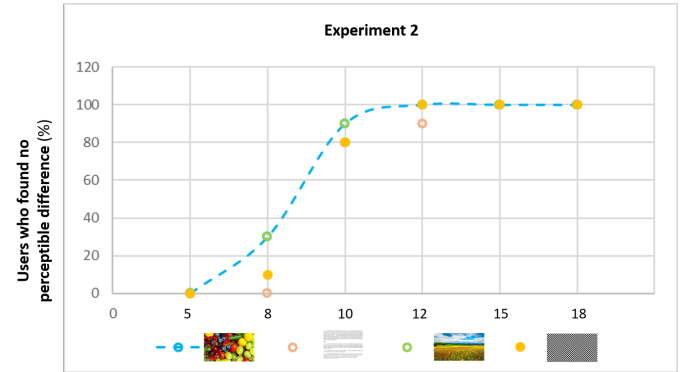Fig. 7.    Percentage of users who did not perceive blur vs. $(\sigma_m, \sigma_m)$



Fig. 8.    Sample image with blurred middle and peripheral regions

## E. Results

In the first experiment, the point at which each participant reported a difference, marked the threshold values of blur for that person. This can be seen in Fig. 7

The threshold for blur was determined approximately based on the percentage of users who didn't perceive that blur. Beyond this, a majority of the participants perceived the blur. Also, it was observed that the 'tunnel vision' was more prominent in the case of the high frequency checkerboard and text images, as compared to natural images, for the same model parameters. Also, higher luminance of content in the peripheral region aided perceptual quality.

In the second experiment, the foveal radius was decreased for the fixed blur values obtained from experiment 1. Different participants started perceiving the blur at different foveal radii. This can be seen in Fig. 8

Thus from these two experiments it was observed that the perception quality varies widely from person to person. However, the best approximate threshold values for $\sigma_m$ and $\sigma_p$ were found to be $0.5$ and $1.8$ respectively and the smallest imperceivable size of foveal region is around $10°$.

## VI. Discussion and Future Work

The technique for foveated rendering explored in this project only points towards potential computational savings. Currently, a real-time foveated rendering system actually suffers an overhead in calculating and implementing the spatial blur. The potential speed-up can be realized by a hardware architecture that utlizes the psychometric response observed in this, and related papers, and renders only amount of pixels corresponding to the magnitude of blur as a function of eccentricity.

Moreover, work in this project is only presented as a prototype and there could be many possible enhancements. A few are mentioned below:

- Implement real-time 3D rendering pipeline with requisite hardware to overcome latency hindrances.
- Experiment for larger FOV setting.
- Extend to a VR HMD display and conduct user study to be able to learn effects coupled with other effects like vergence-accommodation conflict.
- Explore different MAR discretization and blurring models to optimize trade-off between computational efficiency and perceptual quality.

## VII. ACKNOWLEDGEMENTS

## REFERENCES

[1] Understanding foveated rendering — sensics. http://sensics.com/understanding-foveated-rendering/. (Published on 4/11/2016).

[2] Andrew T Duchowski, Nathan Cournia, and Hunter Murphy. Gaze-contingent displays: A review. *CyberPsychology & Behavior*, 7(6):621–634, 2004.

[3] Thomas A Funkhouser and Carlo H Séquin. Adaptive display algorithm for interactive frame rates during visualization of complex virtual environments. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 247–254. ACM, 1993.

[4] Brian Guenter, Mark Finch, Steven Drucker, Desney Tan, and John Snyder. Foveated 3d graphics. *ACM Transactions on Graphics (TOG)*, 31(6):164, 2012.

[5] Eric Horvitz and Jed Lengyel. Perception, attention, and resources: A decision-theoretic approach to graphics rendering. In *Proceedings of the Thirteenth conference on Uncertainty in artificial intelligence*, pages 238–249. Morgan Kaufmann Publishers Inc., 1997.

[6] Anjul Patney, Marco Salvi, Joohwan Kim, Anton Kaplanyan, Chris Wyman, Nir Benty, David Luebke, and Aaron Lefohn. Towards foveated rendering for gaze-tracked virtual reality. *ACM Transactions on Graphics (TOG)*, 35(6):179, 2016.

[7] Hans Strasburger, Ingo Rentschler, and Martin Jüttner. Peripheral vision and pattern recognition: A review. *Journal of vision*, 11(5):13–13, 2011.

[8] Hector Yee, Sumanita Pattanaik, and Donald P Greenberg. Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments. *ACM Transactions on Graphics (TOG)*, 20(1):39–65, 2001.