Exploring Foveation Techniques for Virtual Reality Environments

Razeen Hussain[a], Manuela Chessa[b] and Fabio Solari[c]
Department of Informatics, Bioengineering, Robotics and Systems Engineering, University of Genoa, Genoa, Italy

Keywords: Foveation, Virtual Reality, Immersive Media, Image Quality, Gaze-Contingency, Visual Perception.

Abstract: Virtual reality technology is constantly advancing leading to the creation of novel experiences for the user. High-resolution displays often are accompanied by higher processing power needs. Foveated rendering is a potential solution to circumvent this issue as it can significantly reduce the computational load by rendering only the area where the user is looking with higher detail. In this work, we compare different foveated rendering algorithms in terms of the quality of the final rendered image. The focus of this work is on evaluating 4K images. These algorithms are also compared based on computational models of human visual processing. Our analysis revealed that the non-linear content-aware algorithm performs best.

1 INTRODUCTION

In the realm of virtual reality (VR), the pursuit of unparalleled realism and immersive experiences has driven continual advancements in technology. One of the focal points of development over the years has centered on enhancing display resolution. As VR applications seek to reproduce and surpass real-world visual experiences, the drive for higher display pixel densities proves pivotal in delivering lifelike visuals, sharper details, and a higher sense of presence for users. Nevertheless, certain perceptual challenges persist (Hussain et al., 2023), presenting avenues for further exploration and refinement in the pursuit of an even more compelling VR experience.

A pivotal challenge faced by VR developers lies in optimizing computational resources without compromising visual fidelity. Foveation techniques, rooted in the human visual system's ability to focus sharply on specific regions while perceiving peripheral areas with lower acuity, have emerged as a promising solution. These techniques selectively allocate rendering resources, concentrating detail where the user is looking and reducing the computational load in the peripheral vision (Mohanto et al., 2022). When combined with accurate eye-tracking functionality, foveated rendering could accelerate the creation of large-screen displays with more expansive fields of view and higher pixel densities (Roth et al., 2017). As VR applications become increasingly diverse, from gaming and simulations to medical and educational contexts, understanding the nuances of different foveation methods is crucial.

Foveation can be applied at various stages of the rendering pipeline. They can be applied to optics, screen space, or object space (Jabbireddy et al., 2022). Optics-based foveation modifies the display's optical system using eye tracking to adjust focus, prioritizing high-detail rendering in the user's gaze area. Screen-space foveation optimizes computational performance by rendering central image details and gradually reducing detail towards the periphery. Object-space foveation preprocesses 3D model geometry, employing multiple models with decreasing levels of detail based on the user's gaze, reducing rendered polygons for enhanced computational efficiency.

The aim of this work is to evaluate recent foveated rendering algorithms. In particular, we evaluate how natural the output is since in the quest for immersive experiences, achieving naturalness is paramount; it denotes how seamlessly the foveated rendering technique integrates high-resolution focus areas with lower-resolution peripheral regions, mirroring the way human vision naturally perceives visual stimuli.

In this work, we assess 7 foveation techniques on 4K data, aligning with the growing prevalence of high-resolution VR displays. We compare them in terms of computation time and image quality assessment (IQA), utilizing both referenced (PSNR, SSIM, VIF, FovVDP) and reference-less (BRISQUE, NIQE, and PIQE) metrics. By employing metrics inspired by

[a] https://orcid.org/0000-0002-7579-5069
[b] https://orcid.org/0000-0003-3098-5894
[c] https://orcid.org/0000-0002-8111-0409

Hussain, R., Chessa, M. and Solari, F.
Exploring Foveation Techniques for Virtual Reality Environments.
DOI: 10.5220/0012458200003660

In *Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2024) - Volume 1: GRAPP, HUCAPP and IVAPP*, pages 321-328
ISBN: 978-989-758-679-8; ISSN: 2184-4321

computational models of visual perception, our goal is to determine the compatibility/suitability of VR-based foveated techniques with the human visual system since such techniques are used in VR with humans. This could also serve as a valuable tool for the early testing and assessment of foveation algorithms, minimizing the reliance on experimental sessions with human participants until their final development stage.

2 FOVEATED RENDERING

Foveated rendering is a technique used in VR and other visual systems to optimize the allocation of computational resources and bandwidth (Jin et al., 2020). It leverages the concept of foveation, which mimics the human visual system by concentrating higher detail in the central visual field (foveal region) and reducing detail in the peripheral vision. This means allocating more pixels to represent the central part of an image or video and fewer pixels for the surrounding areas (Hussain et al., 2019). By prioritizing the allocation of resources to the areas where users are most likely to focus their attention, foveation aims to reduce the overall amount of data that needs to be transmitted or processed, thus improving bandwidth efficiency. This approach is particularly relevant in VR applications, where high computational demands and limited bandwidth can impact the quality and responsiveness of the virtual experience.

Spatial resolution adaptation in foveated rendering often involves log-polar mapping, a process where the original image is transformed into both cortical and retinal domains (Solari et al., 2012). This transformation results in an image with higher resolution at the center, gradually decreasing as one moves towards the periphery. Foveated rendering techniques, such as the kernel-based approach (Meng et al., 2018), leverage this spatial variation. In this method, the high-acuity foveal region is synchronized with head movements, while the peripheral region aligns with the virtual world, optimizing computational efficiency.

Some systems have introduced algorithms that compute multiple resolution images, subsequently constructing the final image through a combination of the high and low resolution images for the foveal and the peripheral areas respectively (Romero-Rondón et al., 2018). A challenge of such approaches is the occurrence of artifacts in the transitional regions. A blending function can be incorporated to minimize these artifacts (Hussain et al., 2020).

A prevailing challenge in foveated rendering lies in determining optimal parameters. Traditional methods rely on fixed parameters, but recent advancements, like content-aware prediction model (Tursun et al., 2019), introduce adaptability based on luminance and contrast. A study on foveal region size and its impact on cybersickness revealed that users adapt more swiftly to larger foveal regions (Lin et al., 2020).

However, geometric aliasing remains a persistent issue, manifesting as temporal flickering. Solutions, such as temporal foveation integrated into the rasterization pipeline (Franke et al., 2021) can dynamically decide whether to re-project pixels from the previous frame or redraw them, and is especially effective for dynamic objects. Another approach involves post-processing by adding depth-of-field (DoF) to mitigate artifacts (Weier et al., 2018). While this method showed promising visual results, challenges related to achieving necessary frame rates underscore the importance of combining DoF blur and foveated imaging for optimal outcomes.

Recently, the image quality of foveated compressed videos has been evaluated (Jin et al., 2021). The authors use various objective metrics as well as subjective measures to assess the quality of the images with various compression levels. They focus on live images which are more appropriate for 360°VR. On the other hand, our work focuses on animated data which is more common in traditional VR setups.

3 EXPERIMENTAL STUDY

The aim of this study is to provide valuable insights into the degree of naturalness achieved by different foveation techniques. Since the support resolution of consumer VR devices is gradually increasing, a goal of this work is to evaluate the performance of these techniques on 4K data.

3.1 Foveation Techniques

A spectrum of foveation methods was considered based on the type of algorithm and availability of code. In the end, 7 techniques for foveation or foveated rendering were shortlisted. These range from traditional gaze-based foveation to advanced machine learning-driven approaches. The techniques are briefly described below:

3.1.1 MRF

The multi-region foveation (MRF) process begins with scene analysis to identify the primary area of interest. A Gaussian blur is applied to the peripheral regions of the visual field, gradually attenuating

spatial details. Concurrently, the central region undergoes higher resolution rendering, preserving critical details. The degree of foveation is dynamically adjusted based on user focus, ensuring adaptability to changing visual contexts. The implementation involves a two-step rendering process. First, the low-resolution peripheral image is generated, and then the high-resolution central image is overlaid. Although the image can be divided into circular or rectangular sections (Bastani et al., 2017), we use circular egocentric regions as they better correlate to the optics of the majority of consumer VR devices.

3.1.2 SVIS

Space variant imaging system (SVIS) is a toolbox available for MATLAB (Geisler and Perry, 2008). It allows the simulation of foveated images in real time. The foveation encoder and the foveation decoder are the two components that make up this foveated imaging algorithm. The original image is put through a number of low-pass filtering and downsampling processes in the foveation encoder to produce a pyramid of images with progressively lower resolutions. The low-pass pictures from the pyramid are up-sampled, interpolated, and blended by the foveation decoder to create a displayable image that is smoothly foveated.

3.1.3 LPM

The technique proposed in (Solari et al., 2012) has the aim of mimicking the non-linear space variant sampling of the human retina, i.e. the log-polar mapping (LPM). This technique is well suited to computing visual features directly in the cortical domain since it allows for defining a proper spatial sampling pattern. Moreover, it has a fast implementation since it uses bilinear interpolation instead of a Gaussian blur. This implementation also allows obtaining a spatial sampling closer to the human one that is not completely described by a Gaussian blur. Indeed, this technique is used also to mimic human visual processing, such as the ones of disparity (Maiello et al., 2020) and optic flow (Chessa et al., 2016). It is also possible to exploit the inverse mapping in order to use the technique for foveated rendering. Specifically, we use LPM parameters similar to the ones in (Maiello et al., 2020) for performing the comparison proposed in this work.

3.1.4 Contrast Foveation

This technique introduces a perceptually-based foveated real-time renderer designed to approximate a contrast-preserving filtered image (Patney et al., 2016). The motivation behind this approach is to address the sense of tunnel vision introduced by reduced contrast in the filtered image. To achieve contrast preservation, the renderer pre-filters certain shading attributes while undersampling others. To mitigate temporal aliasing resulting from under-sampling, a multi-scale version of temporal anti-aliasing is applied. Both pre- and post-filtering reduce contrast, which is normalized using a post-process foveated contrast-enhancing filter.

3.1.5 Noised Foveation

This technique proposed by (Tariq et al., 2022) enhances foveated images by considering the spatial frequencies. The authors highlight a limitation in contemporary foveated rendering techniques, which struggle to distinguish between spatial frequencies that must be reproduced accurately and those that can be omitted. The process begins with a foveated image as the input and estimates parameters such as orientation, frequency, and amplitude for Gabor noise. Subsequently, Gabor kernels are generated based on these estimated parameters and convolved with random impulses, effectively synthesizing procedural noise. In the next step, the synthesized procedural noise is introduced to the contrast-enhanced foveated image. This addition of noise contributes to the overall visual richness and complexity of the image, enhancing the perceived details and textures in the foveal region. The integration of Gabor noise in this manner is a deliberate strategy to augment the visual quality and realism of the foveated image.

3.1.6 Aware Foveation

This technique proposes a computational model for luminance contrast to determine the maximum amount of spatial resolution loss that may be added to an image without causing observable artifacts (Tursun et al., 2019). The model incorporates elements such as a contrast perception transducer model and peripheral contrast sensitivity, which are based on aspects of the human visual system. The model predictions are fine-tuned by utilizing acquired experimental data. The predictor model's ability to predict parameters accurately for high-resolution rendering even when it is only given a low-resolution image of the current frame is one of its key primary features. This feature is essential for determining the required quality before producing the entire high-resolution image.

3.1.7 Foveated Depth-of-Field

The foveated depth-of-field is a technique originally designed to mitigate the onset of cybersickness in VR systems (Hussain et al., 2021). The algorithm is implemented in the screen space as a post-processing

effect. It takes into account how the human visual system works in real-world viewing and attempts to minimize the discrepancies between it and virtual object viewing. It radially divides the image into three sections, representing the foveal, near-peripheral, and mid-peripheral regions. The output image is further refined by computing the depth-of-field. Thus, mimicking real-world viewing where objects placed at the accommodative distance regardless of their position in the field-of-view are perceived with high fidelity. It incorporates a blending function to remove artifacts in the transitory regions where there is a significant difference in blurring between adjacent pixels.

3.2 Dataset

We use the MPI Sintel dataset (Butler et al., 2012) which contains 23 scenes capturing diverse indoor and outdoor scenes using Blender. Overall, the dataset contains 1064 8-bit RGB images where each image has a resolution of 1024x436 and contains corresponding depth maps. The choice of the dataset was motivated by the fact that some of the foveation techniques rely on depth maps to refine the output. Since, the data is not of 4K resolution, a pre-processing step was performed to scale the data (see Section 3.4).

3.3 Evaluation Metrics

The objective of this study is to analyze the naturalness of images produced by foveation. For this purpose, we use both referenced and reference-less IQA metrics. Metrics that take into account the human visual system were also included. Overall, the following metrics were selected for comparison:

- **Execution Time.** The execution time and frame rate are critical factors in VR applications due to their direct impact on user experience and immersion. In VR environments, users interact with computer-generated content that must be rendered and updated in real-time to create a seamless and immersive experience. Therefore, a lower processing time is desired.
- **PSNR.** Peak Signal-to-Noise Ratio (PSNR) quantifies the difference between the two images in terms of both signal fidelity and noise introduced during compression. Higher PSNR values indicate better quality.
- **SSIM.** Structural Similarity Index (SSIM) (Wang et al., 2004) is a metric used to quantify the similarity between two images by taking into account luminance, contrast, and structure. The SSIM index ranges from -1 to 1, where higher SSIM values generally correspond to better perceptual image quality.
- **VIF.** Visual Information Fidelity (VIF) (Sheikh and Bovik, 2006) is based on the concept that the human visual system is sensitive to various frequency components in an image. It quantifies how well an image preserves important visual information when compared to a reference image. It considers both luminance adaptation and contrast sensitivity functions of the human visual system. The metric ranges from 0 to 1, where 1 indicates perfect similarity.
- **FovVDP.** FovVDP (Mantiuk et al., 2021) is a full-reference visual quality metric that predicts the perceptual difference between pairs of images and videos. It is aimed at comparing a ground truth reference video against a distorted version, such as a compressed or lower framerate video. FovVDP works for videos in addition to images, accounts for peripheral acuity, and works with SDR and HDR content. It models the response of the human visual system to changes over time as well as across the visual field, so it can predict temporal artifacts like flicker and judder, as well as spatio-temporal artifacts as perceived at different degrees of peripheral vision. Such a metric is important for head-mounted displays as it accounts for both the dynamic content, as well as the large field of view.
- **BRISQUE.** Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) (Mittal et al., 2011) is a no-reference image quality metric that operates on a machine learning algorithm, utilizing natural scene statistics to gauge image quality. The BRISQUE score is determined by comparing the image to a default model generated from images of natural scenes exhibiting comparable distortions. A lower score is indicative of higher perceptual quality.
- **NIQE.** Natural Image Quality Evaluator (NIQE) (Mittal et al., 2012) is a no-reference image quality metric that gauges the naturalness of an image by measuring it against a default model derived from images of natural scenes. A lower NIQE score signifies superior perceptual quality.
- **PIQE.** Perception-based Image Quality Evaluator (PIQE) (Venkatanath et al., 2015) is an image quality evaluator that calculates the no-reference image quality score using a perception-based approach by measuring the local variance of perceptibly distorted blocks. The PIQE score is a non-negative scalar in the range [0, 100]. A higher score value indicates lower perceptual quality.

3.4 Experimental Procedure

The experimentation was performed on a workstation comprised of an Intel Core i7-9700K processor and an NVIDIA GeForce 1080 graphics card. The algorithms were run using MATLAB implementations. We chose MATLAB as our primary platform for implementation because we intend to compare foveated rendering algorithms based on computational models of the human visual system, allowing for a comprehensive analysis within a controlled environment.

Since the dataset is of a lower resolution than 4K, a pre-processing step was performed to scale the images to the desired resolution (3840x2180). In order to maintain the aspect ratio and subsequently not introduce any distortion, the pre-processing was performed in two stages. First, the image was resized using the same aspect ratio as that of the dataset and then, a section of the enlarged image was cropped corresponding to the 4K resolution from the center of the image. A similar operation was also performed on the depth maps.

The center of the image was chosen as the gaze location. This was motivated by the fact that many VR user studies have found that the users' fixation point tends to be on the center of the display the majority of the time during an immersive experience (Clay et al., 2019) while the users prefer to move their head and not just the gaze when they want to fixate on other objects.

The overall process is shown in Figure 1. The original image is first transformed to the desired resolution. Parameters such as gaze location, Gaussian pyramid levels, depth maps, etc. are passed to the foveation algorithm as per requirement. The output image shown in the figure is of Noised Foveation. Evaluation metrics are computed accordingly.

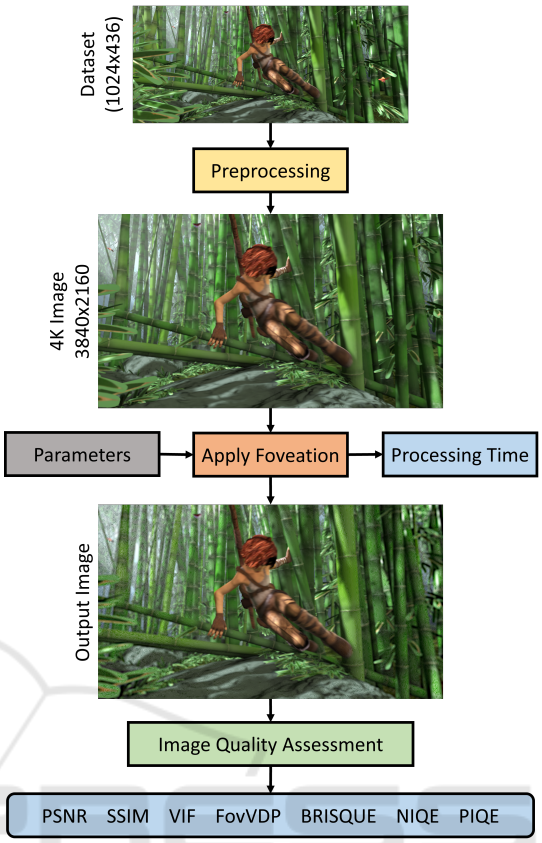

Figure 1: The overall process flow.

4 RESULTS

The experimental analysis is shown in Table 1, taking into account metrics such as computation time and referenced and reference-less IQA. For the reference-less image quality metrics, the corresponding metric score for the original image is also shown to provide a comparison.

Example output images of the foveation techniques are shown in Figure 2. However, the differences may not be apparent. Therefore, maps/masks corresponding to pixels with noticeable differences from the reference image and potential areas of noticeable artifacts are also shown. These include the SSIM maps which highlight local values of SSIM, FovVDP difference maps which contain color-coded visualization of the difference map in the form of heatmaps and PIQE artifacts which highlight the number of artifacts in the image noticeable to the human eye.

In the referenced image quality metrics, the Aware Foveation technique emerged as the top performer. This can be attributed primarily to the model's capability to predict the sigma level individually for each pixel, determining the degree of reduction in spatial resolution. This precision in predicting and adjusting the spatial resolution at the pixel level contributes to achieving superior image quality. The model's performance is notably comparable to the non-linear bio-inspired technique of LPM, further emphasizing its efficacy in capturing and optimizing the intricacies of spatial resolution across the image.

Among the VR-based foveated rendering techniques, Foveated DoF emerged as the second best performing. This was largely due to the artifact reduction step incorporated into the algorithm and to the incorporation of accommodative distance, a property of the lens in the human eyes. Contrast Foveation performed better than the MRF algorithm. It should be noted that the Contrast Foveation used the output of the MRF as

Table 1: Comparison among foveation algorithms. The mean values over the dataset are reported along with the standard deviation in rounded brackets. The best and second best performing algorithm for each metric is highlighted in bold and bold-italics respectively. Metrics for the original data are also reported in the case of reference-less image quality metrics.

	Original	MRF	SVIS	LPM	Contrast Enhanced	Noised Foveation	Aware Foveation	Foveated DoF
Time (s)	–	6.584 (0.090)	***1.237*** ***(0.014)***	1.568 (0.045)	11.726 (0.204)	93.703 (0.816)	72.394 (1.321)	**0.453** **(0.029)**
PSNR	–	35.631 (4.243)	41.953 (3.869)	***48.285*** ***(3.743)***	37.168 (4.276)	28.957 (2.751)	**53.457** **(3.117)**	45.730 (5.180)
SSIM	–	0.984 (0.012)	0.995 (0.004)	***0.998*** ***(0.002)***	0.985 (0.011)	0.880 (0.047)	**0.999** **(0.001)**	***0.998*** ***(0.002)***
VIF	–	0.494 (0.063)	0.697 (0.045)	0.848 (0.032)	0.537 (0.065)	0.423 (0.059)	***0.930*** ***(0.022)***	**0.956** **(0.019)**
FovVDP	–	9.569 (0.186)	9.739 (0.117)	***9.943*** ***(0.044)***	9.724 (0.122)	9.456 (0.117)	**9.983** **(0.014)**	9.329 (0.345)
BRISQUE	54.823 (5.460)	55.469 (9.135)	***52.941*** ***(6.328)***	56.991 (6.294)	55.753 (8.446)	**13.993** **(6.466)**	55.548 (5.871)	56.039 (4.677)
NIQE	5.876 (0.470)	6.031 (0.548)	6.163 (0.605)	5.698 (0.569)	6.100 (0.515)	**2.802** **(0.240)**	5.654 (0.454)	***5.531*** ***(0.550)***
PIQE	77.194 (8.214)	80.942 (9.565)	81.370 (8.898)	***77.520*** ***(8.069)***	81.580 (9.520)	**28.486** **(5.325)**	78.265 (7.889)	79.035 (7.130)

an input and the difference between the two is relatively small.

Although Noised Foveation was the worst performing in the referenced image assessment, it was the best performing in the reference-less metrics. This may be due to the fact that the considered reference-less metrics take into account the difference in texture across the image. The artifacts introduced by the foveation technique, largely due to the incorporation of noise into the algorithm, may have influenced the metric scores. However, it should be noted that the number of noticeable artifacts is quite high (see PIQE artifacts in Figure 2). These are located mostly in the peripheral regions and may not be an issue since some VR user studies have found that the human visual system is more sensitive to artifacts present inside 20°of eccentricity (Hoffman et al., 2018).

Nevertheless, the discrepancy in the values of reference-less metrics, notably lower than those of the original images, raises concerns about the potential influence of noise and suggests a need for caution in interpreting and reporting these metrics. It is possible that these metrics are susceptible to misinterpretation due to noise. This issue will be thoroughly addressed in future work through experimental sessions involving participants to ensure a more accurate and reliable evaluation.

In terms of computation time, the Foveated DoF algorithm performed best, followed by SVIS and LPM. Although we report the computation time each foveated imaging algorithm took, it should be noted that some of the algorithms under evaluation have been designed for implementation on OpenGL-enabled GPUs, and as a consequence, their performance experiences a notable slowdown when executed using MATLAB on a CPU. It is imperative to recognize that these algorithms are inherently optimized for GPU architecture, and their efficiency is expected to significantly improve when deployed in actual VR applications, where computations take place on a dedicated GPU using shaders, ensuring a notably faster execution.

Overall, the Aware Foveation model exhibited the highest performance, closely followed by Foveated DoF and LPM. We opted to assign higher weightage to the FovVDP metric since it accounts for the physical display specification (size, viewing distance), foveated viewing, and temporal aspects of vision.

5 CONCLUSIONS

The primary contribution of this study lies in conducting a systematic evaluation of various VR-based foveated rendering algorithms and their comparison through computational models of human visual perception. A comprehensive assessment was performed on 7 algorithms, each applicable to the integration of foveation within VR applications. The evaluation encompassed key criteria such as accuracy, the presence of artifacts, and computation time.

The findings from this research underscore the critical importance of considering the targeted hardware environment and the specific metrics employed

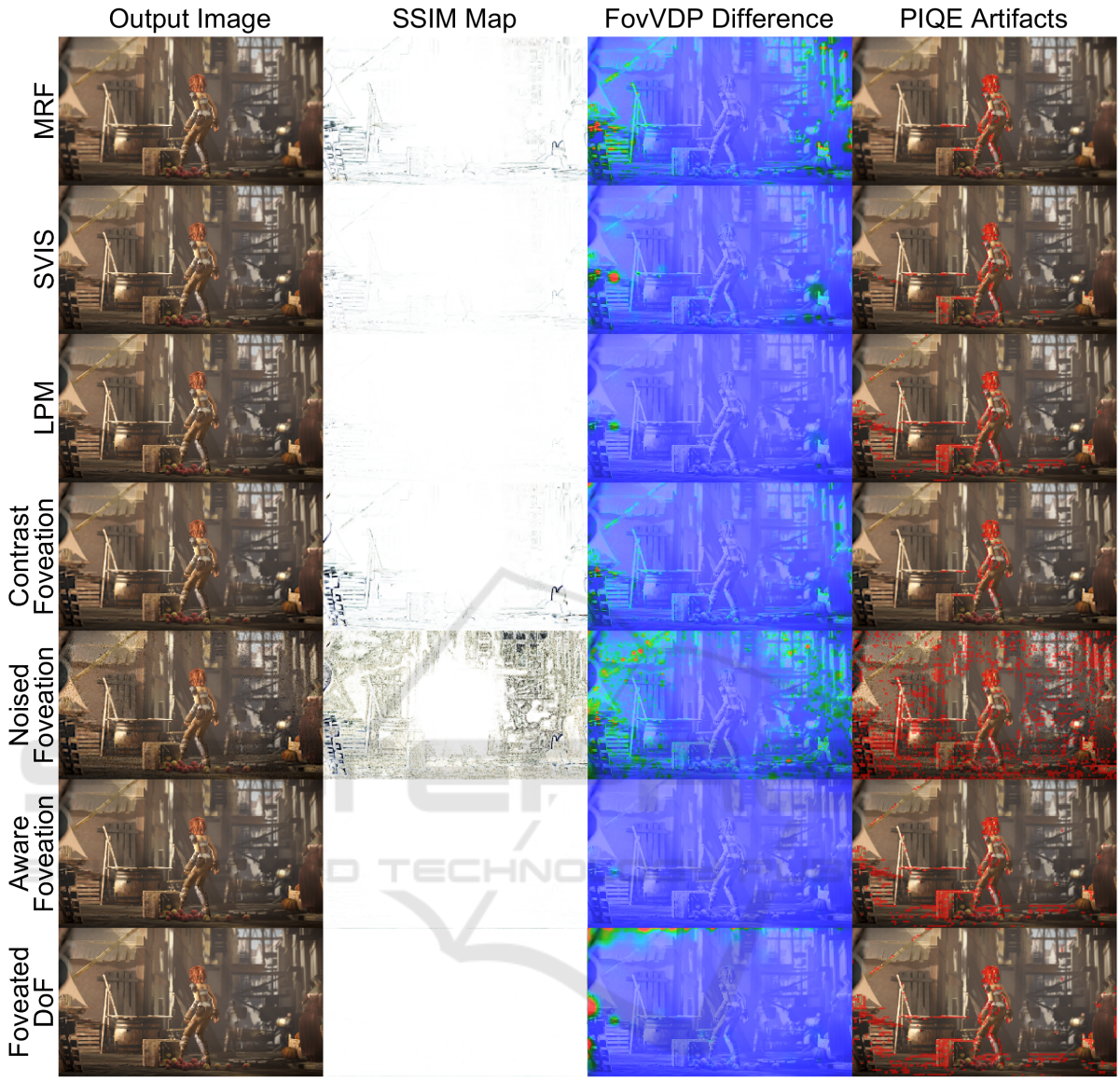

Figure 2: The output images of the tested foveation techniques. Maps highlighting differences or artifacts are also shown.

when assessing and selecting foveated imaging algorithms for practical implementation in the immersive domain of virtual reality. This recognition is vital for tailoring the choice of foveation techniques to the unique demands and constraints of diverse hardware configurations and application scenarios. As a future work, we intend to extend this investigation by conducting a subjective study. This user study will aim to compare the outputs of the considered foveated rendering techniques from a perceptual standpoint, providing a more holistic understanding of their performance and impact on the user experience.

REFERENCES

Bastani, B., Turner, E., Vieri, C., Jiang, H., Funt, B., and Balram, N. (2017). Foveated pipeline for ar/vr head-mounted displays. *Information Display*, 33(6):14–35.

Butler, D. J., Wulff, J., Stanley, G. B., and Black, M. J. (2012). A naturalistic open source movie for optical flow evaluation. In A. Fitzgibbon et al. (Eds.), editor, *European Conference on Computer Vision (ECCV)*, Part IV, LNCS 7577, pages 611–625. Springer-Verlag.

Chessa, M., Maiello, G., Bex, P. J., and Solari, F. (2016). A space-variant model for motion interpretation across the visual field. *Journal of Vision*, 16(2):1–24.

Clay, V., König, P., and Koenig, S. (2019). Eye tracking in virtual reality. *Journal of Eye Movement Research*, 12(1).

Franke, L., Fink, L., Martschinke, J., Selgrad, K., and Stamminger, M. (2021). Time-warped foveated rendering for virtual reality headsets. *Computer Graphics Forum*, 40(1):110–123.

Geisler, W. S. and Perry, J. S. (2008). Space Variant Imaging System (SVIS). https://svi.cps.utexas.edu/svistoolbox-1.0.5.zip.

Hoffman, D., Meraz, Z., and Turner, E. (2018). Limits of peripheral acuity and implications for vr system design. *Journal of the Society for Information Display*, 26(8):483–495.

Hussain, R., Chessa, M., and Solari, F. (2020). Modelling foveated depth-of-field blur for improving depth perception in virtual reality. In *4th IEEE International Conference on Image Processing, Applications and Systems*, pages 71–76.

Hussain, R., Chessa, M., and Solari, F. (2021). Mitigating cybersickness in virtual reality systems through foveated depth-of-field blur. *Sensors*, 21(12).

Hussain, R., Chessa, M., and Solari, F. (2023). Improving depth perception in immersive media devices by addressing vergence-accommodation conflict. *IEEE Transactions on Visualization and Computer Graphics*, pages 1–13.

Hussain, R., Solari, F., and Chessa, M. (2019). Simulated foveated depth-of-field blur for virtual reality systems. In *16th ACM SIGGRAPH European Conference on Visual Media Production*, London, United Kingdom.

Jabbireddy, S., Sun, X., Meng, X., and Varshney, A. (2022). Foveated rendering: Motivation, taxonomy, and research directions. *arXiv preprint arXiv:2205.04529*.

Jin, Y., Chen, M., Bell, T. G., Wan, Z., and Bovik, A. (2020). Study of 2D foveated video quality in virtual reality. In Tescher, A. G. and Ebrahimi, T., editors, *Applications of Digital Image Processing XLIII*, volume 11510, page 1151007. International Society for Optics and Photonics, SPIE.

Jin, Y., Chen, M., Goodall, T., Patney, A., and Bovik, A. C. (2021). Subjective and objective quality assessment of 2d and 3d foveated video compression in virtual reality. *IEEE Transactions on Image Processing*, 30:5905–5919.

Lin, Y.-X., Venkatakrishnan, R., Venkatakrishnan, R., Ebrahimi, E., Lin, W.-C., and Babu, S. V. (2020). How the presence and size of static peripheral blur affects cybersickness in virtual reality. *ACM Transactions on Applied Perception*, 17(4):1–18.

Maiello, G., Chessa, M., Bex, P. J., and Solari, F. (2020). Near-optimal combination of disparity across a log-polar scaled visual field. *PLOS Computational Biology*, 16(4):1–28.

Mantiuk, R. K., Denes, G., Chapiro, A., Kaplanyan, A., Rufo, G., Bachy, R., Lian, T., and Patney, A. (2021). Fovvideovdp: A visible difference predictor for wide field-of-view video. *ACM Transactions on Graphics*, 40(4).

Meng, X., Du, R., Zwicker, M., and Varshney, A. (2018). Kernel foveated rendering. *Proceedings of ACM on Computer Graphics and Interactive Techniques*, 1(1).

Mittal, A., Moorthy, A. K., and Bovik, A. C. (2011). Blind/referenceless image spatial quality evaluator. In *45th ASILOMAR Conference on Signals, Systems and Computers*, pages 723–727. IEEE.

Mittal, A., Soundararajan, R., and Bovik, A. C. (2012). Making a "completely blind" image quality analyzer. *IEEE Signal Processing Letters*, 20(3):209–212.

Mohanto, B., Islam, A. T., Gobbetti, E., and Staadt, O. (2022). An integrative view of foveated rendering. *Computers & Graphics*, 102:474–501.

Patney, A., Salvi, M., Kim, J., Kaplanyan, A., Wyman, C., Benty, N., Luebke, D., and Lefohn, A. (2016). Towards foveated rendering for gaze-tracked virtual reality. *ACM Transactions on Graphics*, 35(6).

Romero-Rondón, M. F., Sassatelli, L., Precioso, F., and Aparicio-Pardo, R. (2018). Foveated streaming of virtual reality videos. In *9th ACM Multimedia Systems Conference*, MMSys '18, page 494–497, New York, NY, USA. Association for Computing Machinery.

Roth, T., Weier, M., Hinkenjann, A., Li, Y., and Slusallek, P. (2017). A quality-centered analysis of eye tracking data in foveated rendering. *Journal of Eye Movement Research*, 10(5).

Sheikh, H. and Bovik, A. (2006). Image information and visual quality. *IEEE Transactions on Image Processing*, 15(2):430–444.

Solari, F., Chessa, M., and Sabatini, S. P. (2012). Design strategies for direct multi-scale and multi-orientation feature extraction in the log-polar domain. *Pattern Recognition Letters*, 33(1):41–51.

Tariq, T., Tursun, C., and Didyk, P. (2022). Noise-based enhancement for foveated rendering. *ACM Transactions on Graphics*, 41(4).

Tursun, O. T., Arabadzhiyska-Koleva, E., Wernikowski, M., Mantiuk, R., Seidel, H.-P., Myszkowski, K., and Didyk, P. (2019). Luminance-contrast-aware foveated rendering. *ACM Transactions on Graphics*, 38(4).

Venkatanath, N., Praneeth, D., Maruthi Chandrasekhar, B., Channappayya, S. S., and Medasani, S. S. (2015). Blind image quality evaluation using perception based features. In *21st National Conference on Communications*, pages 1–6.

Wang, Z., Bovik, A., Sheikh, H., and Simoncelli, E. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612.

Weier, M., Roth, T., Hinkenjann, A., and Slusallek, P. (2018). Foveated depth-of-field filtering in head-mounted displays. *ACM Transactions on Applied Perception*, 15(4):1–14.