Slide1

Starting from now, I will begin the presentation by Nozaki from the Zhang Laboratory.　The research theme is the examination of foveated imaging processing based on visual field characteristics.

Slide2

There's an increasing demand for remote operation of equipment situated in distant areas, such as long-distance drone control and remote operation in hazardous disaster sites.

Slide3

In these remote operation systems, enhancing the quality of transmitted images is crucial for improving the operator's sense of presence. However, existing systems often sacrifice image quality uniformly to conserve network bandwidth, by methods like downsizing images or reducing frame rates.

Slide4

Transmission bit rate refers to the number of bits transmitted per unit time in digital data transmission. While higher transmission bit rates generally lead to better video quality, they also demand more network bandwidth. This poses challenges in scenarios with limited network bandwidth, such as remote locations.

Slide5

To tackle these challenges, we introduce Foveated Imaging processing, which selectively allocates resolution based on the human visual field's characteristics. By displaying high-resolution images in the fovea and lower-resolution images in peripheral areas, we can reduce the transmission bit rate while maintaining the appearance of high resolution.

Slide6

The central visual field can be divided into two regions: the discrimination field and the effective field. In graph (a), the discrimination field refers to the central foveal region where visual functions such as acuity are most excellent. In graph (b), the effective field refers to the area between the discrimination field and the peripheral field. Moreover, recent studies have shown that within the same eccentricity, the visual acuity is higher along the horizontal axis compared to the vertical axis, and furthermore, the visual acuity is superior in the lower visual field compared to the upper visual field.

Slide7

Most current Foveated Imaging processing assumes uniform visual acuity horizontally and vertically and processes images in circular regions. Considering human visual field characteristics, the discriminative field is circular, and the effective field is elliptical. Unnecessary areas are also processed with high resolution in image processing.

Slide8

My study addresses the limitations of existing foveated imaging processing by incorporating circular discrimination fields and elliptical effective fields, aiming to reduce transmission bit rates while preserving visual quality.

Slide9

Explanation of the principle of foveated imaging processing: In step 1, a mask image is generated. In step 2, a Gaussian blurred image is created from the original image. In step 3, the original image and the blurred image are overlaid according to the mask image to create a foveated image. The Gaussian filter can remove high-frequency components of the image, enabling reduction of the transmission bitrate.

Slide10

We divided the effective visual field region into layers and created mask images: Circular Mask, Elliptical Mask, and Elliptical Mask based on visual field characteristics. We validated four types of foveated images with different central fovea regions.

Slide11

We used drone footage captured in Full HD resolution for verification. Samples 2 and 3 contained significant high-frequency components, making them high-bitrate videos.

Slide12

The experiment includes verifying visual quality equality, validating transmission bit rate reduction, and verifying processing speed.

Slide13

In Experiment 1, we conducted a kernel size identification experiment using the DCR method to evaluate image degradation with different Gaussian filter kernel sizes.

Slide14

The DCR (Double Stimulus Continuous Quality Scale Rating) method is a subjective quality assessment technique used for evaluating video quality. This method compares the target video with a reference video, as shown in the diagram, to assess the degree of degradation.

Slide15

The evaluators will be asked to assess the degree of degradation using a 5-point scale, and based on the results, the optimal kernel size will be identified. In this evaluation process, the circular mask model will be utilized. Evaluators will be instructed to fixate on the center of the screen and evaluate whether there is overall degradation in the video compared to the reference video.

Slide16

We had 10 students from the laboratory evaluate the videos. Table 2 shows the average results for each kernel size. According to the DCR method, a mean value of 3.5 is considered the "threshold of acceptability," and videos with kernel sizes of 13 and 15 exceeded this threshold. Table 3 presents the average results for each video when the kernel size was 13. Since all videos had average values exceeding 3.5, we set the parameter to 13 and proceeded with further experiments.

Slide17

In Experiment 2, we will use the DCR method to verify whether there is no difference in visual quality between circular and elliptical regions. We will use circular model videos for the reference video and compare it with videos using elliptical models and elliptical models based on visual field characteristics for target vide.

Slide18

In Experiment 2, the average value obtained for the elliptical region was 4.83, while the average value for the elliptical region based on visual field characteristics was 4.78. Both values exceeded the "detection threshold" of 4.5, suggesting that visual quality is maintained compared to the circular region.

Slide19

In Experiment 3, we will compare the transmission bit rates among four conditions: full resolution, circular, elliptical, and elliptical based on visual field characteristics.

Slide20

Here are the results for Experiment 3:

Table 5 shows the transmission bit rates and reduction rates for Sample 1. Comparing circular and elliptical based on visual field characteristics, a 3.6% improvement in reduction rate was achieved.

Table 6 shows the transmission bit rates and reduction rates for Sample 2. Comparing circular and elliptical based on visual field characteristics, a 3.3% improvement in reduction rate was achieved.

Slide21

Table 7 shows the transmission bit rates and reduction rates for Sample 3. Comparing circular and elliptical based on visual field characteristics, a 7.4% improvement in reduction rate was achieved.

Table 8 shows the transmission bit rates and reduction rates for Sample 4. Comparing circular and elliptical based on visual field characteristics, a 1.0% improvement in reduction rate was achieved.

Based on these results, it can be concluded that using elliptical regions based on visual field characteristics rather than circular regions leads to improved reduction in transmission bit rates.

Slide22

In Experiment 4, we will compare the processing time of the foveated imaging processing among three models. The experiments were conducted in an environment with a CPU and memory of 16GB.

Slide23

The experimental results are as follows:

Comparing processing times, the elliptical model increased by 1.7 times compared to the circular model, while the elliptical model based on visual field characteristics increased by 2.87 times. Furthermore, it was observed that the processing time increases with larger video data sizes. Most of the processing time was attributed to the generation of mask images.

Slide24

From these results, it can be concluded that using an elliptical model based on visual field characteristics for foveated imaging processing maintains visual quality.

Additionally, transitioning from a circular to an elliptical model based on visual field characteristics allows for a reduction in transmission bit rate by 1% to 7.4% compared to circular models. Despite an increase in processing time by 2.87 times compared to circular models.

It is evident that while processing time increases, it is still possible to reduce transmission bit rates while maintaining visual quality.

Slide25

Future tasks include speeding up processing time and developing real-time systems by incorporating viewer gaze information.

Slide22

Thank you all for your attention.