Usage Report for Fiscal Year 2023

Project Title:
Development of human visual system-inspired image processing system
Name:

Name:
Name: ○ Zhe Sun
Laboratory at RIKEN:
(1) Center for Advanced Photonics, Image Processing Research Team

1. Background and purpose of the project, relationship of the project with other projects

Human vision characteristics play a vital role in achieving perceptually appealing compression outcomes, leading to a growing interest in integrating vision models into image compression. Nevertheless, much of the prior research has focused on direct compression of original images, without incorporating any scene understanding. The human eye's anatomy features a sparse and biased distribution of photoreceptors that capture visual information. In our project, we are developing an image compression framework guided by importance, grounded in a model of human vision.

2. Specific usage status of the system and calculation method

Our framework aims to incorporate image understanding into the image encoding and decoding tasks. For any given input image, our encoding process selectively samples salient pixels based on an importance map inspired by human vision. During decoding, the raster image is reconstructed from the encoded pixels using models based on prior information. Presently, we employ tensor completion for image reconstruction in the decoding process, which is time-intensive. We are exploring parallel algorithms with HOKUSAI for tensor completion to address this issue.

3. Result

A system was created that integrates dithering processing, targeted at areas of interest in input images for compression, with restoration using tensor completion methods. This technique enabled compression of images to 1% of the original pixel count based on visual attention. However, compressing to less than 0.03% of the original pixel count proved challenging, and the time required for restoration was substantial. In response, the research shifted towards developing methods based on Generative Adversarial Networks (GAN) and diffusion model. Additionally, new saliency-weighted evaluation metrics were devised. These metrics, when applied in the proposed method, appear to effectively evaluate the quality of images in areas of interest, thus making them appropriate for assessing perceptual characteristics.

4. Conclusion

We have introduced an image compression framework that leverages a human visual importance map. This new framework efficiently encodes an input image by adaptively sampling color pixels from the estimated importance map, aligning well with the prior information-based models used in our decoding process. This approach effectively retains the visual features of the image. Our numerical experiments showed that the proposed framework integrates visual attention and geometric features into the compression process more effectively than traditional sampling methods.

Furthermore, our results suggest that enhancing the CNN-based restoration models could lead to promising applications in retinal model simulations and other image processing tasks using our vision-inspired importance maps.

5. Schedule and prospect for the future

Currently, we are exploring learning-based methods, such as generative adversarial networks and diffusion models, for image reconstruction. In the future, we also plan to modify our algorithm to make it applicable to video.

6. If no job was executed, specify the reason.