DISSERTATION ABSTRACT

Generation of images for stereoscopic displays using selected perceptual features of human visual system

mgr inż. Marek Wernikowski
supervisor: dr hab. inż. Radosław Mantiuk, prof. ZUT

The dissertation deals with the generation and visualization of three-dimensional images on stereoscopic displays using binocular vision. The stereoscopic effect is created in these displays by displaying independent images for the right and left eyes. This technology allows for a correct and realistic perception of depth in the 3D scene. On the other hand, rendering two or more images becomes more challenging in computations. The dissertation proposes solutions that reduce the sampling of the scene and thus accelerate the synthesis of images. The solution above uses the features of the human visual system to ignore information that is irrelevant from the point of view of human perception. In particular, the characteristics of non-uniform resolution of vision (visual acuity) and variable sensitivity to contrast visual crowding are used.

The content of the work describes five related methods that, using the perceptual features of the human visual system (HVS), accelerate the real-time synthesis and display of images.

The foveated rendering system is based on visual acuity changes along the visual field, particularly its reduction in peripheral vision. Thus, the number of samples can be reduced for areas far away from the gaze direction measured by the eye tracker. The sampling can also be changed depending on the scene's content because content that is invisible to humans can be omitted or rendered with reduced accuracy. For this purpose, a directional contrast sensitivity model has been developed. The proposed foveated rendering method improves the shading operation speed more than twice compared the state-of-the-art foveated rendering systems.

An extension of the above method is the proposed anti-aliasing technique, which detects temporal aliasing in peripheral vision, where the sampling frequency is significantly reduced. The method is based on a convolutional neural network trained using animated scenes that have been generated using the directional rendering method.

The phenomenon of visual crowding was used to replace image fragments with metamers, which are different from the original image, but not in a way detectable by a human. Especially in the peripheral areas of vision, the metamers can be fundamentally different from images synthesized by classical techniques, making it possible to limit the number of samples used to generate peripheral image areas. The dissertation examines the possibility of using metamers for directional rendering. Metamers are created using GAN training (generative adversarial network). A novelty of the solution is the use of distorted images to train the network, which finally allows for a lower sampling rate of the scene.

A complex image decomposition into two image planes is necessary for stereoscopic multifocal displays. The dissertation proposes a hybrid decomposition method based on a perceptual model of contrast sensitivity. It uses a rigorous decomposition method only for the image's content, which may be distorted by a quick interpolation method. The developed method was tested in the custom multifocal display.

The dissertation also introduced a tone mapping technique that combines the foveated rendering width visual adaptation phenomenon. The adaptation luminance is computed based on the surrounding brightness of the gaze point. In this way, the effect of maladaptation, i.e., the smooth change of adaptation to brightness over time, is simulated.

10.05.22
Marek Wernikowski