PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

An application of the virtual prototyping approach to design of VR, AR, and MR devices free from the vergence-accommodation conflict

Igor S. Potemin, Irina Livshits, Dmitry Zhdanov, Andrey Zhdanov, Nikolay Bogdanov

Igor S. Potemin, Irina Livshits, Dmitry Zhdanov, Andrey Zhdanov, Nikolay Bogdanov, "An application of the virtual prototyping approach to design of VR, AR, and MR devices free from the vergence-accommodation conflict," Proc. SPIE 10694, Computational Optics II, 1069404 (28 May 2018); doi: 10.1117/12.2314095

SPIE.

Event: SPIE Optical Systems Design, 2018, Frankfurt, Germany

An application of the virtual prototyping approach to design of VR, AR, and MR devices free from the vergence-accommodation conflict

Igor S. Potemin*, Irina Livshits, Dmitry Zhdanov, Andrey Zhdanov, Nikolay Bogdanov
ITMO University, 49 Kronverksky Pr., St. Petersburg, 197101, Russia

ABSTRACT

One of the main problems in the design of AR, VR and MR devices is an estimation of the design result. Base optical characteristics like aberrations or contrast functions are not applicable. On the other hand, the real prototyping of such devices is very expensive. As a result, authors propose to use a virtual prototyping approach as an element of design of AR, VR and MR devices that allows avoiding real prototyping and therefore to reduce design cost. The virtual prototyping approach assembles all elements of the VR, AR or MR system from the system of the virtual image generation to the eye perception. This approach was implemented as a virtual parametric model of that device.

Keywords: Virtual prototyping, vergence-accommodation conflict, virtual reality, augmented reality, mixed reality, eye tracking, optical design, image processing

1. INTRODUCTION

Systems of virtual and augmented reality have found wide application in many areas of human activity. In particular they are head-up and head mounted in the aircraft building and automotive industry, virtual or augmented reality headsets in virtual prototyping systems, training systems and simulators when creating intuitive interfaces for interacting with objects of the surrounding world in systems of augmented and mixed realities.

The high quality of optics, the realism of the virtual world image and the high resolution of the micro-display are not able to provide a natural perception of the virtual world. For the naturalness of perception, it is necessary to ensure the reconciliation of vergence and accommodation of the human eyes. The greatest problems in this regard may arise in the stereoscopic systems for video 360[1,2]. In these systems, the shooting of one part of the scene is carried out by a pair of co-directional cameras, displaced relative to each other by the average value of the center-to-center distance of the human eyes. This corresponds to a practically infinite focusing distance when observing the environment. When viewing such a video, the observer can comfortably perceive objects that are at practical infinity. If an object in the field of view is found near the camera, then when you try to focus at this object, the image will be doubled, because the direction of the survey corresponded to the infinitely distant object. This problem cannot be eliminated within the framework of viewing video 360 and therefore a possible recommendation how to reduce the impact of this problem is to organize the survey itself without a close-up.

In VR systems, this problem can be partially solved. When the direction of view is changed, i.e. when you focus on the selected object of the virtual world, you must immediately synthesize two new images of the virtual scene that correspond to the direction of the view of the right and left eyes. Naturally, for the synthesis of new images it is necessary to define a new direction of view. This direction is determined using special eye tracking devices. However, the correct orientation of the cameras corresponding to the viewing direction of the left and right eyes used in the synthesis of the real world images is not yet a sufficient condition for the formation of a natural perception of the virtual world image. A sufficient condition is the consistency of accommodation and vergence of human eyes[3,4,6]. Therefore, when a person looks at an object, they not only converge the directions of the eyes of the left and right eyes on this object, but also focuses on it, i.e. compresses or relaxes the lens of the eye, in order to see this object as clearly as possible (while the remaining objects are blurred).

The blur effect associated with the accommodation of the eyes is due to the finite size of the pupil of the eye. From the optical point of view, the eye is a photographic lens with a variable focal length. It can focus on certain objects, but because of the finite size of the pupil, objects outside the focusing plane are blurred.

*ipotemin@yandex.ru; phone 7 921 930-0821; fax 7 812 457-1537

Computational Optics II, edited by Daniel G. Smith, Frank Wyrowski, Andreas Erdmann, Proc. of SPIE
Vol. 10694, 1069404 · © 2018 SPIE · CCC code: 0277-786X/18/$18 · doi: 10.1117/12.2314095

The blur size depends on the amount of the object's displacement from the focusing plane (defocusing) and the size of the pupil. The more defocusing and the larger pupil size results in more blur. Computer graphics systems that synthesize an image of the virtual world usually use the model of an ideal point camera with an infinitely small pupil size, as a result of which they are free from the effects of defocusing and blurring of an image. If the computer graphics system synthesizes an image focused on a certain viewing plane, then the image of objects in this plane will be sharp.

In addition, in order to see this image sharp, it is necessary to perform the corresponding diopter adjustment of the eyepiece, i.e. coordinate the position of the eyepiece with the current accommodation of the eye. In this case, the image in the viewing plane will be sharp and look natural. However, the image of all other objects outside this plane will also be clear. And it will look unnatural and will cause discomfort. Figure 1 illustrates this effect.

On the other hand, the absence of an eyepiece adjustment will result in defocusing of the entire image, which will also cause discomfort in the visual perception of the virtual world. Therefore, the problem of reconciling the effects of vergence and accommodation lies in the fact that the formation of an image that is sharp in the plane of convergence of the human eyes and is naturally defocused (in accordance with the eye's accommodation) outside of this plane. Figure 2 illustrates the process of reconciling vergence and accommodation in the formation of images for the left and right eyes.

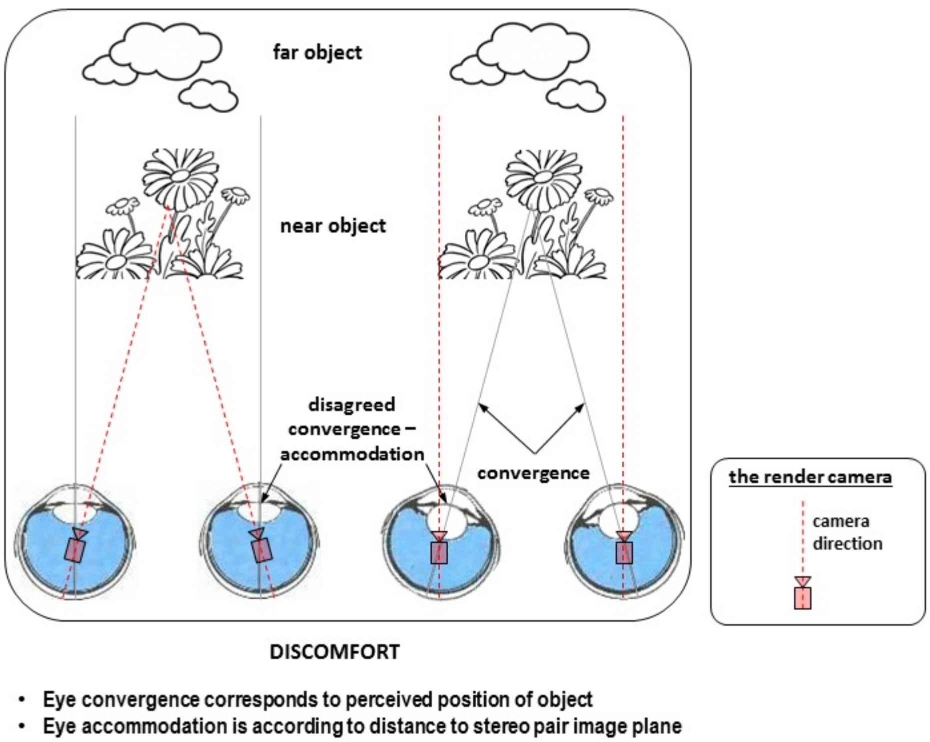

Figure 1. The effect of vergence - accommodation conflict.

To solve the problem of reconciling of the vergence accommodation conflict in the virtual world image, there are several ways[4,5]. The two main ways are either forming a correctly defocused image for an eyepiece with an adaptive focal length or converting a flat image on a microdisplay into a pseudo-volume with a depth corresponding to the depth of the scene.

In the first case, the effect of defocusing caused by the natural accommodation of the human eye is simulated. The defocusing effect is a first-order aberration and the function of scattering a point in the form of a cylinder whose radius is proportional to the defocus value. This defocusing effect can be easily reproduced in any computer graphics system. However, for the natural visualization of a defocused image, the eyepiece should focus on the observation plane in the space of the virtual scene. The eyepiece focusing system must be adaptive and directly related to the pupil tracking system. Changes of the direction of view should be monitored by the pupil tracking system to automatically focus the eyepiece on a new viewing plane. In this case, images for the right and left eyes are automatically rearranged in accordance with the direction of observation and are focused on the current viewing plane. As a result, the discomfort of perception of the virtual world disappears. The only complication of the proposed solution is that the implementation of the adaptive focusing of the eyepiece represents a serious scientific and technical problem that has not yet been solved.

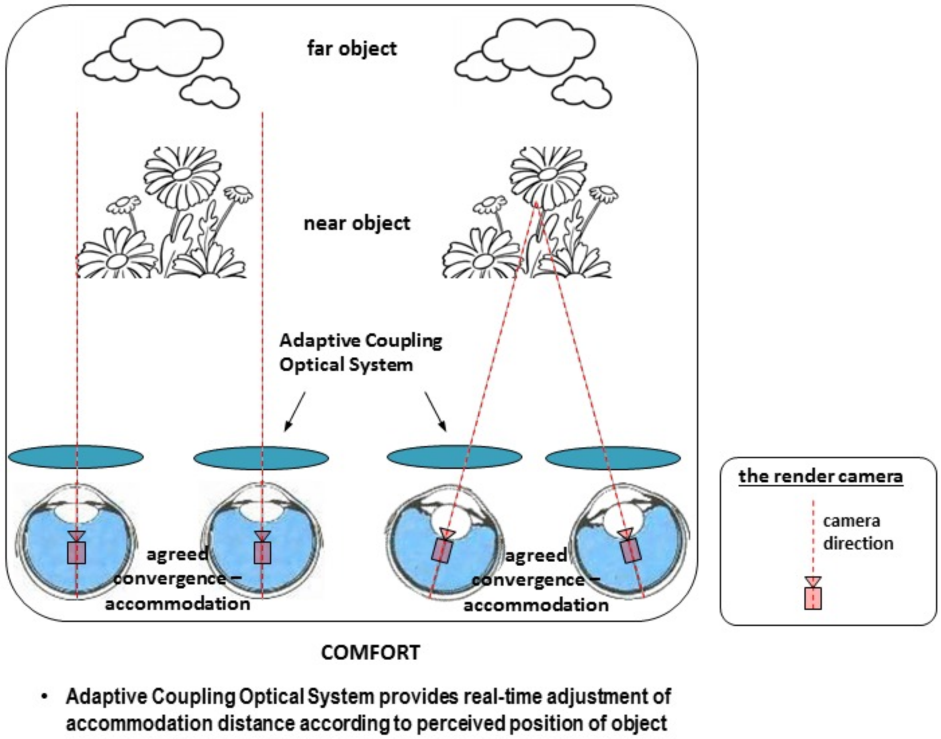

Figure 2. Correct combination of vergence and accommodation in the formation of images for the left and right eyes.

2. RENDERING

The image, which the rendering system forms on the LCD matrix of the virtual reality system, has to be agreed, at first, with an optical system, which transfer the image to the observer eye, and, secondly, with the observation condition (vergence and accommodation), that is what object the person observes. Of course, optical system cannot provide aberration free image on overseer eye and main reason of the aberrations is in hard requirements to the external system parameters (large view field and eye box). Fortunately, the eye pupil is much smaller than the eye box of the optical system and the main contribution to the aberrations is caused by distortion (both chromatic and monochromatic). The modern rendering systems can consider the distortion effects and render image with opposite distortion that finally allows removing the distortion effect from the observed objects of the virtual reality. The remaining problem is agreement between accommodation and distortion of the observed objects of virtual reality. The modern virtual reality devices can be accompanied with an eye tracking system that allow detecting position of the eye pupil and predicting the eye view direction. By processing of the eye tracking information from two eyes, the eye tracking system can predict point in 3D scene where the person eyes are directed. If the person sees an image of the virtual reality then the prediction of the observation position can be more accurate and agreed with the actual scene object, which the person observes.

The image generation with correct parameters of the eye vergence is not a problem for modern rendering system and requires only the correct orientation of two cameras (for left and right eyes). Main problem is correct accommodation. When eye is focused on the object, all other scene objects, which do not belong to the focusing plane, are defocused. The effect of correct focusing (defocusing) can be achieved in two ways. The first way is to transform plane image from the CCD matrix to a pseudo 3D image. According to the calculated distance to the object point (calculated somehow for the corresponding image points), the spatial light modulator can provide phase retardation of the individual pixels of the image and therefore transform 2D image to 2.5D image[5]. As a result, the eye will see an image with correct vergence and accommodation. With exception of the distance map this solution does not requires special extension from the rendering system side and all problems are moved to the optic design side. The map of the distances is not a problem for rendering system and is usually kept in both trace based and raster based rendering systems.

Another way is rendering a 2D image with defocusing corresponding to the real eye perception of the focused distance. The defocusing is a main aberration and it is much higher than all other eye pupil aberrations. Fortunately, in the first order, the modern rendering system can easy add the defocusing aberration to the rendered image. The point spread function (PSP) of the point out of the object plane is cylinder, where radius of the cylinder depends on the focus of the optical system, the pupil diameter and shift of the viewpoint from the object plane. Taking into account that pupil diameter and lens focus are constant for the whole image the radius of PSP is scaled with distance to the point only. As a result, if we have an “ideal” pinhole image and know the distance from the eye pupil to the scene points corresponding to the image point, we can simply convolve the image with the cylindrical function. The only technical difficulty of the convolution is variable radius. However, it is not a problem for modern GPU systems.

Another problem is the physical correctness of this approach. At first, it is a distance to the object. We have to determine distance to the luminance source, not to the first visible scene object. That is if we see some bright object in the mirror we have to apply defocusing procedure to the object visible in the mirror (and calculate distance to the object), not to the mirror. Another problem is effects on the boundaries of two objects placed on the different distances. Taking into account, that pinhole camera does not support the real eye pupil size, the visibility near the object boundary can be different for pinhole and real eye camera. All these effects can be somehow taken into account in the modern systems of computer graphics, and, in most of the real cases, the disagreement will become invisible.

Our main interest was in the area of an analysis and an estimation of the possible disagreement between accommodation and vergence in the systems of virtual, augmented and mixed realities (VR, AR and MR), so we use physically accurate methods of the image rendering (more accurate than methods used in VR, AR and MR systems). Our methods are based on physically correct methods of light propagation, implemented in Lumicept[7] software package that allows creating both realistic scene images for both eyes and rendering images which person really see in the different kinds of VR, AR and MR systems. Comparing images in VR, AR and MR systems with expected images, we can estimate possible discomfort of the eye perception.

3. MODEL OF VERGENCE - ACCOMODATION CONFLICT

The idea of the project is providing the fast and accurate way of prototyping the different ideas and solutions aimed to reduce discomfort of the visual perception caused by disagreement between the accommodation and vergence of human eyes. Currently there are three main approaches dealing with the visualization of the virtual objects. They are VR, AR and MR. Each approach has some specifics, which has to be properly reflected in solutions aimed to the discomfort reduction. In the case of VR all our world is virtual one and we see only images of the virtual objects generated with the rendering system. All objects exist in the rendering system and if VR system is accompanied with the eye tracking system then all information is accessible for the synthesizing of the physically correct rendering image, which takes into account the specificity of the image visualization (how the vergence - accommodation conflict is solved in the optical system). In the case of MR system, the problem is more complex. There are two spaces: real world and virtual world. If rendering system knows all about the virtual world then the real world can be undefined (or not wholly defined) for the rendering system. This ambiguity does not allow to identify the point of the eye convergence on the object of the real world because of it is absent in the rendering system. Therefore, it can result in an additional disagreement between accommodation and vergence of the human eye. AR systems are simplified case of the MR system. Main difference is that virtual objects are mainly in the form of symbols or graphical information, which is placed to the fixed distance (usually infinity distance). Moreover, in most cases the problem of the vergence - accommodation discomfort does not appear because of these systems are not stereoscopic.

Taking into account all problems and complexity of the prototyping of the VR, AR and MR system we propose the specific system of the virtual prototyping. The system consists of six main elements:

- the virtual prototype of the human eye (how image is formed on the eye retina),
- the virtual prototype of the optical system which projects image from the LCD matrix to the eye pupil,
- the virtual prototype of the real world (applicable for AR and MR systems), corresponding to the image which is created on the eye retina when we observe the real world,

- the virtual prototype of the addition optical system on the path from the world to eyes (applicable for AR and MR systems), for example, glasses,
- the virtual prototype of the virtual world, corresponding to the image which is created on the eye retina when we observe objects of the virtual world,
- the virtual prototype of the rendering system, which generate image on the LCD matrix.

As a result, our task looks like:

- creation of the virtual device of VR (AR or MC) from the virtual components listed above,
- generation of the real scene image (of the real or virtual world) on the eye retina (it is supposed that the image is free from the vergence - accommodation conflict),
- generation of the scene images which optical system of VR, AR or MR project from LCD matrix through the optical system to the eye retina,
- comparison of the images (the first is free from the vergence - accommodation conflict and the second is what we really see in VR, AR or MR system).

In the current article we represent results of the virtual prototyping of the VR system with vergence - accommodation conflict and free from the conflict. We analyzed the sources of the conflict and proposed different solutions how to reduce and eliminate the conflict. To eliminate the vergence - accommodation conflict we designed the adaptive eyepiece system with variable focus. The focus is adjusted according the eye convergence (that is the point of the observation) and the rendering system generates images, which corresponds to the images on the eye retinas when eyes are focused on the distance.

Below are the results of our prototyping.

4. SIMULATION RESULTS

To analyze the possible discomfort caused be the vergence - accommodation conflict in the VR system we created a simple scene of the virtual world. Figure 3 illustrates the scene. The observer is placed near the floor so that he sees a number of objects on different distances while all these objects are placed in the view cone of 10 degrees. The fist object is placed on 0.5 meters from the eyes while the last object is on about 32 meters. As a result, the disagreement between vergence and accommodation can be very high and can cause the strong discomfort of the visual perception.

We used two models of the observation. The first model is a model of the natural vision. To create an image on the eye retina we used the computer system of the photorealistic rendering[7]. Images for the left and right eyes were generated separately taking into account all eye effects. That is for each point of the eye convergence we generated the physically correct images with the proper eye accommodation. It means that the image points of all object points placed in the plane of the eye convergence were focused while all other object points were defocused.

The second model of the observation is an observation of the scene of the virtual world in the VR system. We tested two models of the VR system, shown at Figure 4.

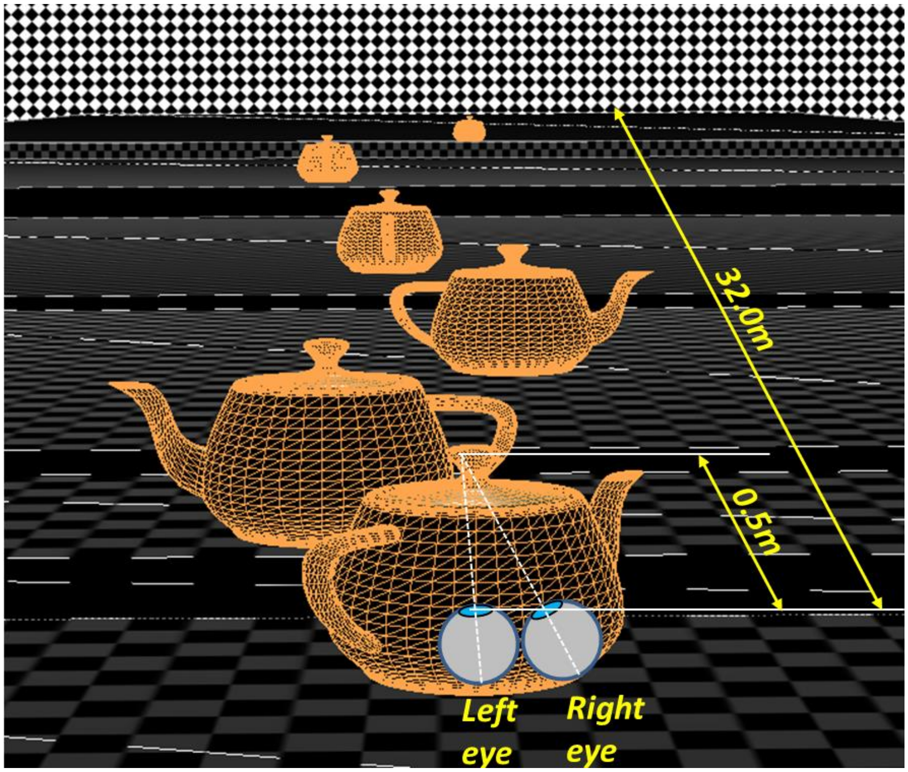

Figure 3. The virtual world scene

The first model of the VR optical system (eyepiece) is a pair of ideal lenses, which provide the aberration free images (Figure 4 shows only one image channel). The model was used as a reference to estimate the quality of the specially designed eyepiece system. Depending on the point of the observation, the LCD is moved, and the eyepiece changes the focus. The second model is a specially designed eyepiece. The quality of the image within a viewfield of 10 degrees is high and this eyepiece can be considered as an aberration free lens too.

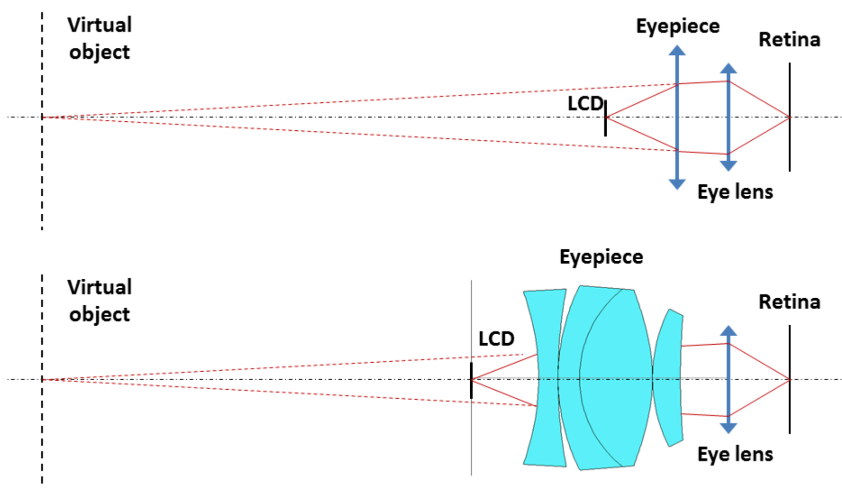

Figure 4. VR system models

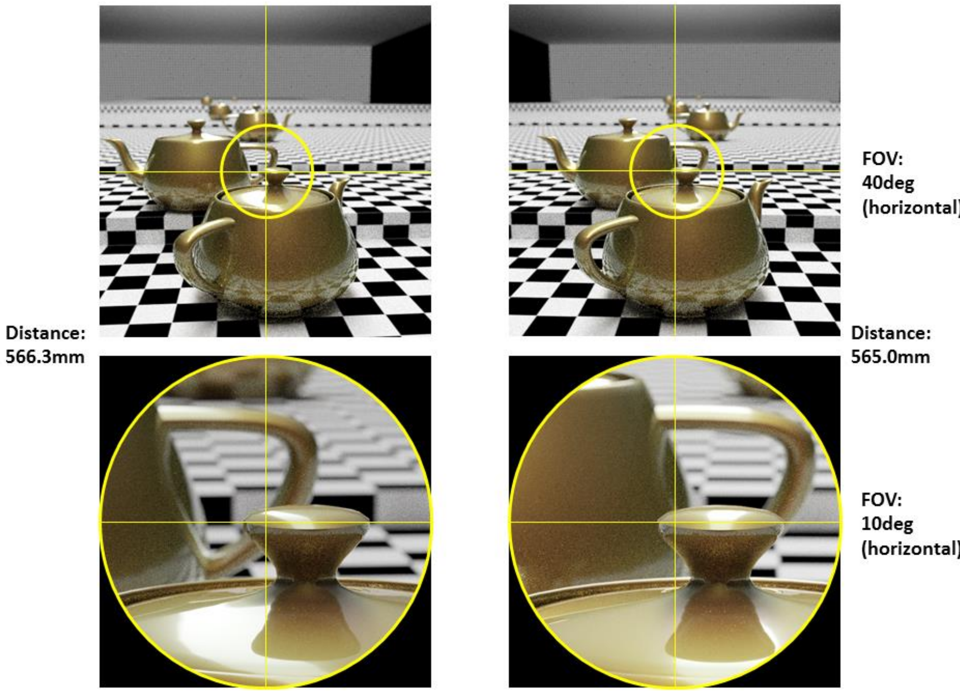

Figure 5. The real view of the VR scene

Figure 5 shows the simulation results. These are images formed on the left and right eyes in the natural conditions. That is what person can see if the scene was the real one. The top part of the image corresponds to the view angle = 40 degrees. The bottom part of the image is an image of the central zone = 10 degree. The human eyes were accommodated and converged to 565 mm. We see the point of the observation in the focus and points out of the focus as defocused.

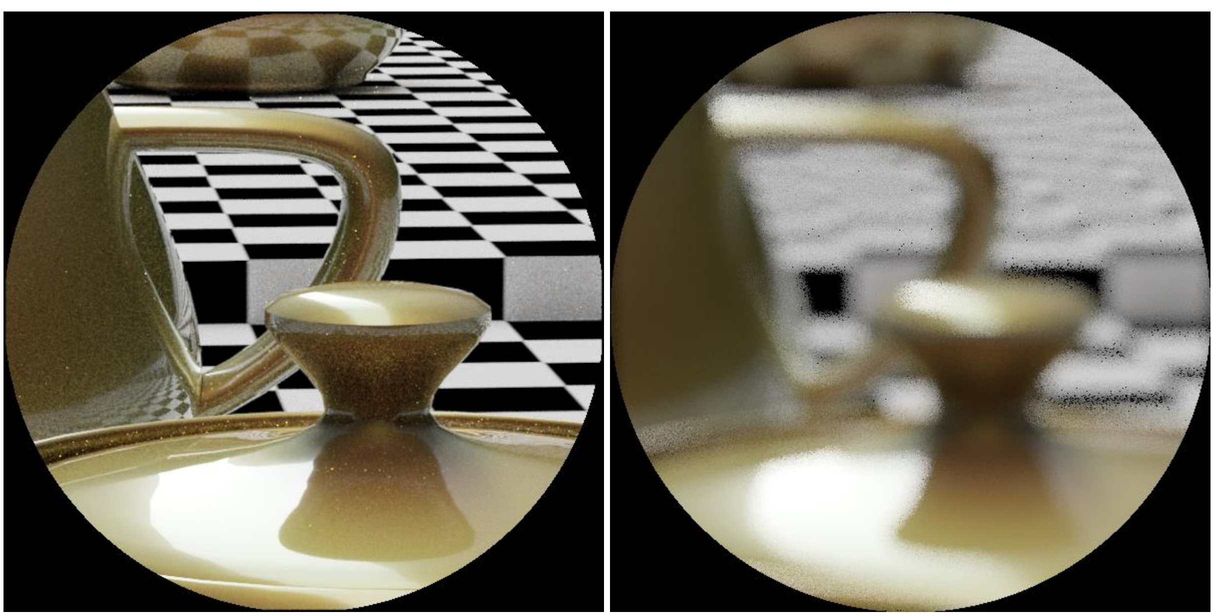

Figure 6. The image viewing in the VR system (the left eye) when LCD is placed in the eyepiece focus, the eye accommodated on infinity (the left image) and accommodation on 565mm (the right image). The rendering uses the pinhole camera

Figure 6 shows the possible conflict between the accommodation and vergence when the rendering system tracks the eye position and provides the correct eye convergence. Two situations are considered and in both cases the rendering system generates an image for the pinhole camera. Moreover, the position of the LCD, placed in the eyepiece focus, is not changed. In the first case the observer is accommodated on the infinity distance. As a result, all objects are in the focus that results in the conflict of the vergence and accommodation. In the second case the observer is accommodated on the distance of 565 mm and sees all objects as defocused ones. The reason is a constant position of the LCD and again we see the conflict of the vergence and accommodation.

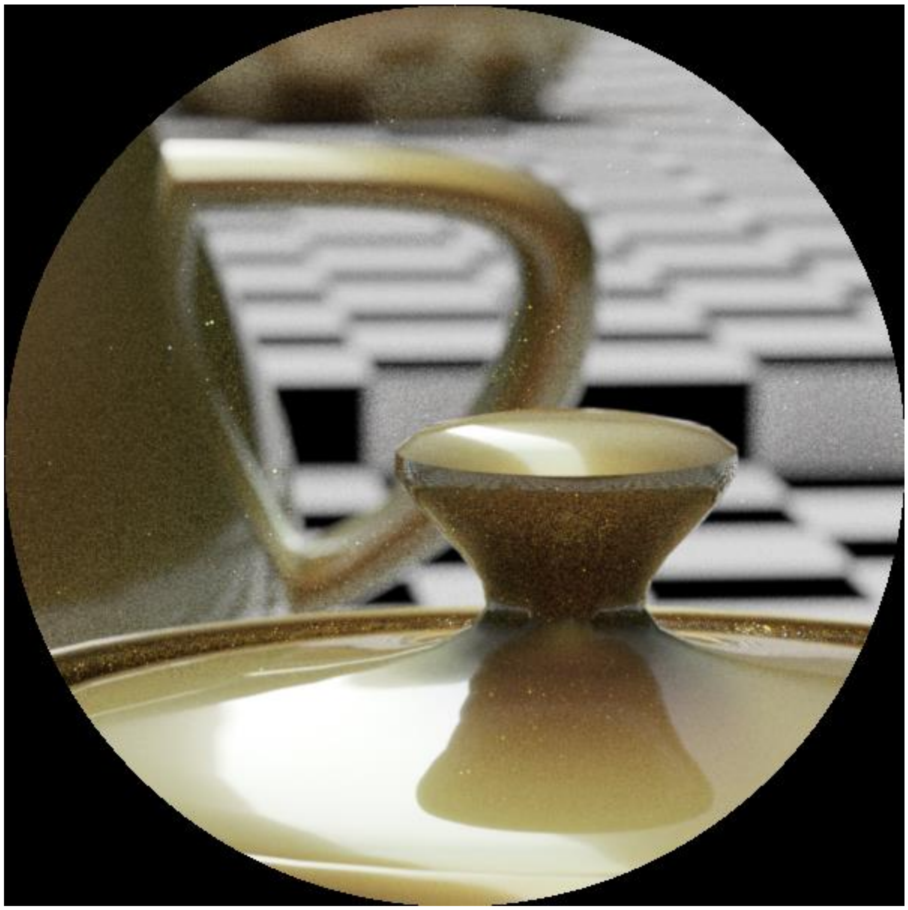

Figure 7. The image viewing in VR system (the left eye) with the shifted LCD and the corrected focal distance of the eyepiece. The eye is accommodated on 565mm

To avoid the conflict, we have to render the image for the proper parameters of the accommodation (565 mm in our case) and modify the optical system. The modification of the optical system includes the modification of the eyepiece focus and the proper shift of the LCD. Figure 7 illustrates the image visible in the VR system when the render image on LCD, the position of the LCD and the eyepiece focus were calculated from the condition of the natural image perception, that is no the vergence - accommodation conflict.

CONCLUSION

As the result of the current research, the solution of the virtual prototyping of the VR, AR and MR systems was proposed. This solution allows estimating and visualizing the vergence - accommodation conflict in these systems and can be used in the design of the VR, AR and MR systems, which are free from the conflict. Example of the VR system, which is free from the conflict, was demonstrated. The continuation of the research will be aimed to the problem of the formalization of the estimation criterion of the conflict value from the result of the virtual prototyping of the VR, AR and MR systems.

ACKNOWLEDGMENTS

The research was partially supported by state financial support of RFBR grants No. 16-01-00552 and 18-08-01484.

REFERENCES

[1] LaValle, S.M., "VIRTUAL REALITY," University of Illinois, Cambridge University Press (2017).
[2] Wetzstein, G., "A personalized VR/AR system that adapts to the user is crucial to deliver the best possible experience," The BRIDGE, Vol. 46, No. 4, Winter (2016).
[3] Peli, E., "Optometric and perceptual issues with head-mounted displays. In Visual Instrumentation," Optical Design and Principles, P. Mouroulis (Ed.). McGraw-Hill, 205-275 (1999).
[4] Kramida, G., "Resolving the Vergence-Accommodation Conflict in Head-Mounted Displays." IEEE TVCG 22, 7, 1912–1931 (2016).
[5] Matsuda, N., Fix, A. and Lanman, D., "Focal Surface Displays," ACM Trans. Graph. 36, 4, Article 86, July, 14 pages (2017).
[6] Robert Konrad, Emily A. Cooper, and Gordon Wetzstein, "Novel Optical Configurations for Virtual Reality: Evaluating User Preference and Performance with Focus-tunable and Monovision Near-eye Displays." ACM CHI (2016), 1211–1220.
[7] "Lumicept – Hybrid Light Simulation Software," Integra Inc., 2018, <http://www.integra.jp/en/products/lumicept> (9 April 2018).