Stereo Presentations Problems of Textual information on an Autostereoscopic Monitor

S.V. Andreev[1,A], N.A. Bondareva[2,B], E.Yu. Denisov[3,A]

[A] Keldysh Institute of Applied Mathematics RAS
[B] Russian Technological University MIREA

[1] ORCID: 0000-0001-8029-1124, esa@keldysh.ru
[2] ORCID: 0000-0002-7586-903X, niki.99@mail.ru
[3] ORCID: 0000-0002-0614-9100, ed@spp.keldysh.ru

Abstract

This paper presents practical experience in constructing stereo presentations of texts and formulas on an autostereoscopic monitor in stereo presentations designed to display the results of numerical simulation. The task of constructing stereo images of texts and formulas is a structural subtask of a general study devoted to the development of methods and algorithms for constructing stereo presentations of the results of scientific research. Previous research has been done for classic stereo devices.This paper discusses the construction of stereoscopic images on an autostereoscopic monitor. The autostereoscopic monitor allows one to observe a stereo image without glasses, while ensuring the quality of the stereo image, which is not inferior to the quality of the stereo image, presented using a classic 3D projection stereo system. Various methods of obtaining stereo images supported by the monitor were tested, namely, the multi-view presentation of the object and the construction of depth maps. The results for both methods are presented. The aim of the study is to create a technology that allows one to quickly and efficiently build full-fledged stereo presentations of scientific results.

Keywords: stereo image, textual information, autostereoscopic monitor, multi-view presentation, depth maps.

1. Introduction

Nowadays, using of stereoanimation to display the obtained scientific results begins to play an important role in a number of industries, in fundamental and applied scientific research.

This is primarily due to the fact that very often a situation arises when a flat image of an object does not provide sufficient information to realize research goals and limits the analysis to the image schematics or the reduced viewing angle.

The stereo image allows to estimate with greater accuracy the relative position of objects in space and, if necessary, fill in the missing data in a given structure beyond the human's gaze, for example, when modeling structural chemical compounds.

Stereo image, in contrast to the two-dimensional flat image, allows one to more effectively explore and simulate objects with a developed spatial structure or interpret photographs taken during experiments with different points in time and integrated into animation.

A lot of articles are devoted quite to this topic. They consider methods for constructing images for various types of stereo devices of both passive and active types. The difference between these types is that the passive ones allow demonstrating this or that material according to a predetermined scenario, however, the viewer has no opportunity to influence the process of the show.

Active (or interactive) installations imply that the viewer has the opportunity to work with the demo material and influence the demonstration process directly. General aspects of imaging for various types of stereo are presented in [1].

Specific problems arising from using a computer system to generate and visualize a composite multi-screen stereo frame. Methods for solving such problems are described in sufficient detail in [1-3].

The gradual spread of modern stereos has generated great interest in the development of practical algorithms for the stereo presentation of static and animated images, which is reflected in papers [4 - 9]. These papers represent the tasks of constructing stereo images in various fields of research: from the Supernova explosion to the process of displacing oil from porous media.

The proliferation of stereo animations and virtual reality also made it possible to reveal a rather important private structural subtask. The standard procedure of a report, presentation of something or a virtual walk implies the presence in them not only of the visual type of information, but also of symbolic - letters, numbers, symbols, metric values, names, numbers, state standards, without which the viewer's perception and assimilation of information is significantly complicated.

 There is accurate information that the person's vision adapts to the stereo image in about 30 seconds. The same time is needed to adapt the vision back to mono image. This makes it impossible to quickly switch the viewer's attention from the stereo image to the usual flat text. In other words, all additional information to the stereo image itself, expressed by graphic symbols (letters, signs or numbers) must also be presented in three-dimensional stereo space, without subjecting the human brain to additional loads.

However, while for geometry and fields of physical quantities, the tasks of constructing stereo representations are developed in sufficient detail, quite a few works are devoted to the development of specific practical approaches and algorithms for constructing stereo images of texts and formulas with sufficient expressiveness and the necessary stereo effect.

Here we can mention only the work [10] describing the study of the perceptibility of stereo fonts on the screens of stereoscopic mobile devices in Japan and our previous works [11,12] on the presentation of textual information for the classic type of stereo. By classic stereo, we mean a 3D projection stereo system for displaying stereo presentations, that includes a graphics station and a projection subsystem of two projectors with a specialized screen.

For the classic type, a linear stereo base was used. The parameters that provide the best effect were found. The main requirements for the used fonts and a number of conditions, the satisfaction of which is necessary to achieve an optimal result, were also revealed [11,12].

In the studies presented in [11,12], the construction of test labels with variations of the font, background, degree of relief of the font, the rotation angle of the caption, shift along the linear stereo base were carried out. Samples of stereo frames were checked on the stereo system in order to evaluate the stereo effect.

Due to the inability to organize a full-fledged expert assessment, the selection of test results was based on the initially chosen principles and the assessment of visual perception of the distance of the test inscription from the screen plane.

Empirically, the value of the shift along the linear stereo base was found, in which viewers located 4 meters from the screen perceived the inscription as being approximately 1 meter away from the screen plane, which was considered sufficient. It was found that the most optimal way to represent the inscriptions is the representation based on the relief font Arial Narrow on a light background, shaded at the bottom, without turning the inscription. An example of a sample recognized as the most successful in all parameters during the course of work is presented in Figure 1.
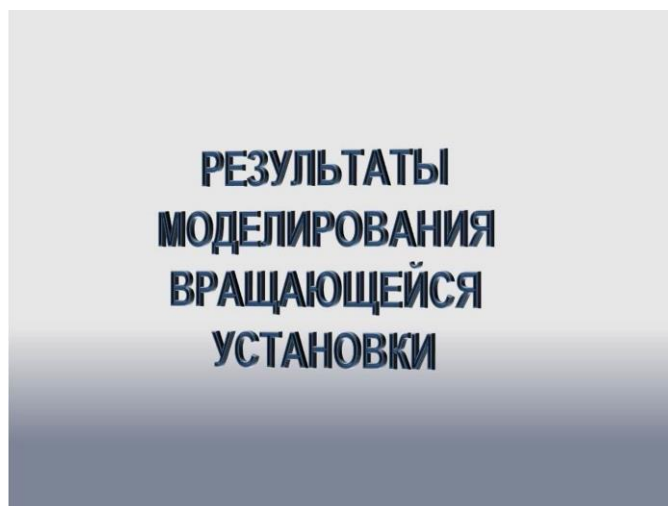
Fig. 1. An example of the construction of the inscription on the basis of the relief font Arial Narrow with the rotation of the inscription for the classic stereo.

This paper continues the cycle of research devoted to the creation of stereo representations, both static and animated, the results of scientific research. The research is carried out in Keldysh Institute of Applied Mathematics (KIAM RAS) on the basis of two available types of stereo devices. The first type of device is a 3D projection stereo system for displaying stereo presentations, educational applications, graphics and films. It is a model of a classic stereographic system using two projectors, a screen and linear polarization glasses

The second type of device is a Dimenco DM654MAS autostereoscopic monitor. Autostereoscopic monitors provide display of stereo images without the need to adjust the equipment depending on the position of the observer. The viewer can move from one segment to another, getting the opportunity to view the displayed object in 3D from different angles of view. The principle of operation of an autostereoscopic monitor is the use of parallax partitions or Fresnel lenses installed behind a protective glass screen. One of the most important advantages of autostereoscopic monitors is the fact that viewing images on such devices does not require the use of special glasses.

The research in KIAM RAS were performed to achieve the most effective result – getting of volumetric stereo images – and to study the properties of the devices themselves to find the most effective algorithms for constructing stereo images for both types of devices. Different methods of construction were used: finding a linear stereo base, depth maps, and some others. A more detailed presentation of the solution of these problems is described in [1-3,7,11,12]. Below are some examples of stereo images for various applications.

Figure 2 shows the results of modeling and visualization of an object depending on various lighting conditions of optically complex materials, such as complex automotive paints.

Fig. 2. Visualization of calculations for optically complex automotive paints on the screen of an autostereoscopic monitor.

Figure 3 presents on an autostereoscopic monitor the results of modeling a supersonic flow around a cone at an angle of attack.
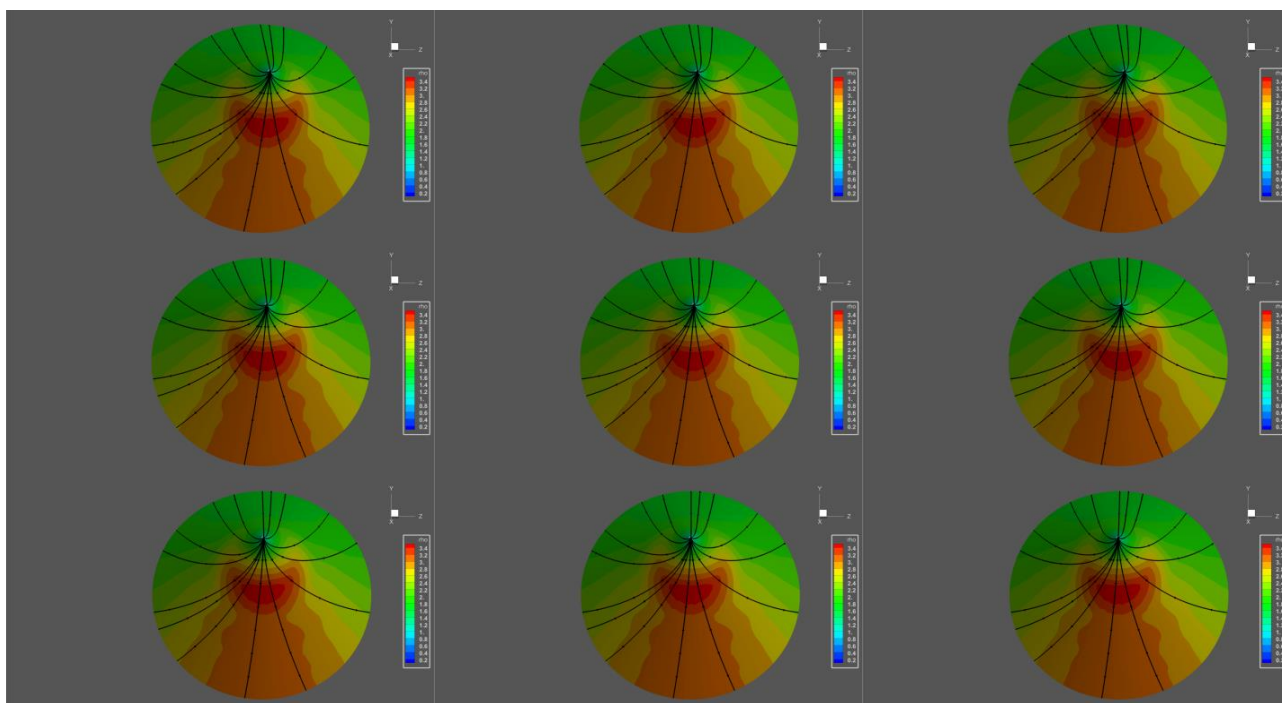

Fig. 3. Visualization the results of modeling a supersonic flow around a cone at an angle of attack for an autostereoscopic monitor.

Figure 4 presents the results of mathematical modeling of a parametric problem for a classical stereo device.
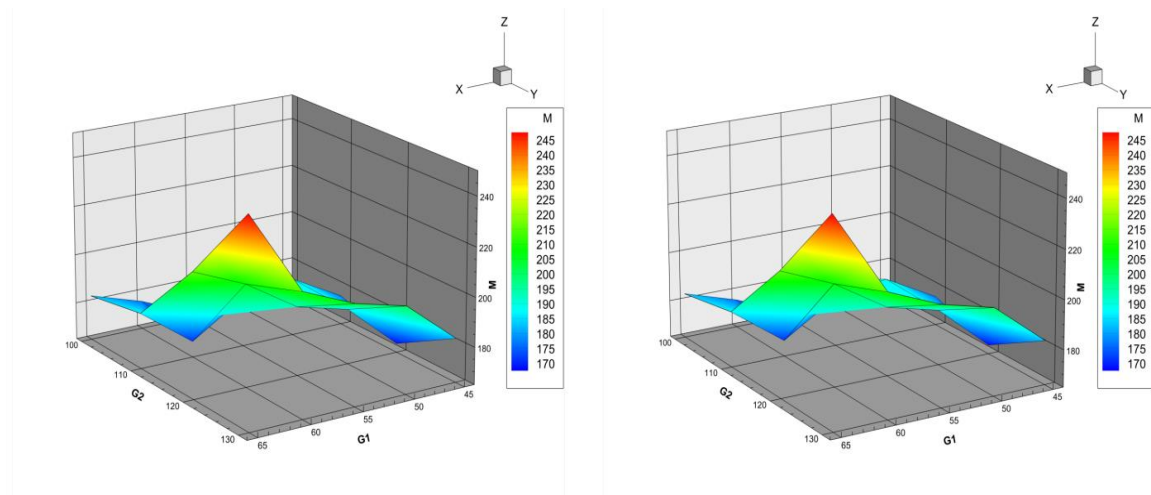
Fig. 4. The results of of mathematical modeling of a parametric problem for a classical stereo device.

At this stage of research, the results of stereo images constructing on an autostereoscopic monitor are presented. The results are obtained by two basic methods of construction: a multi-view presentation and a depth map.

2. Multi-view presentation

An autostereoscopic monitor has the ability to demonstrate an object of visualization using a composite frame containing views of the object of visualization from different angles, which form a certain viewing sector. Usually, these are nine views.

Moreover, these nine views form eight stereo pairs ([1 | 2], [2 | 3], [3 | 4], …, [8 | 9]), and the viewer can watch only one of the stereo pairs, depending on its position in this or that angular sector of observation. Moving from sector to sector, the observer receives volumetric information about the object using all nine angles, that is, as if looking at the object of visualization from different sides.

At this stage of research, a combination of linear and angular displacement was used for the first time. This made it possible to achieve a volume effect comparable to the effect achieved on a classical stereo set.

Nine frames are combined into one image according to the principle of a 3 × 3 matrix (Fig. 5). In the first and last (ninth) frames the inscription is in its extreme positions. The construction of the multi-view representation was organized as follows. In the first frame the inscription is in extreme right position, then with a certain equal intervals it shifts to the left, until it reaches the leftmost position. At the same time, in each frame, the inscription accordingly rotates to the left by the same angle. The axis of rotation passes through the center of the entire inscription, so the leftmost letters are rotated stronger, which creates the effect of immersing them in the screen.
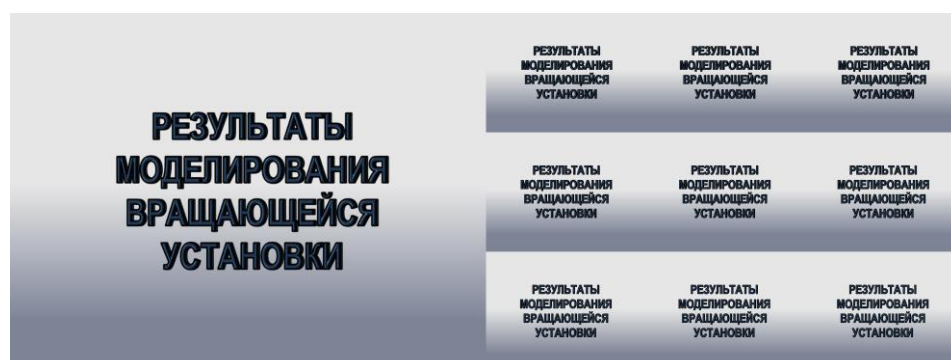

Fig. 5. Constructing a stereo image of an inscription using a multi-view representation.

The task is to ensure that the viewer does not see the transition between each of the stereo pairs, when he changes the viewing angle, so that the strongest realistic effect is created. When the linear and angular shift parameters are too large, artifacts arise: the viewer has a feeling that the image in front of him blurs on transitions from one stereo pair to another, and at the slightest change in position, the viewer notices glare and the boundaries of each of the stereo pairs. Figure 6 shows an unsuccessful pattern using the example of a single letter, where each of the indicated problems stands out particularly brightly: too much rotation of the letter itself and too much linear and angular shift distance.
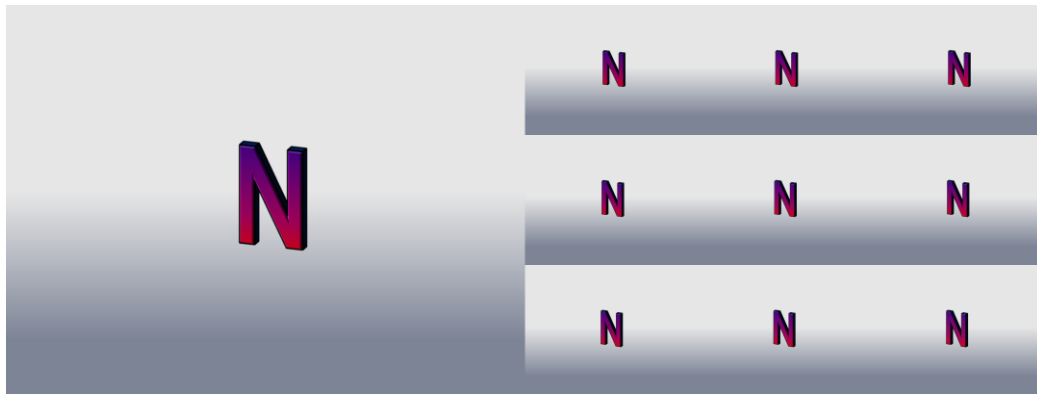

Fig. 6. Constructing a stereo image of a single letter using the multi-view representation.

Experimentally, it was found that if we combine frames in reverse order, when the first and ninth frames, second and eighth, third and seventh, and so on are swapped, then the opposite effect can be achieved, when the inscription "floats" deep into the screen (Fig. 7).
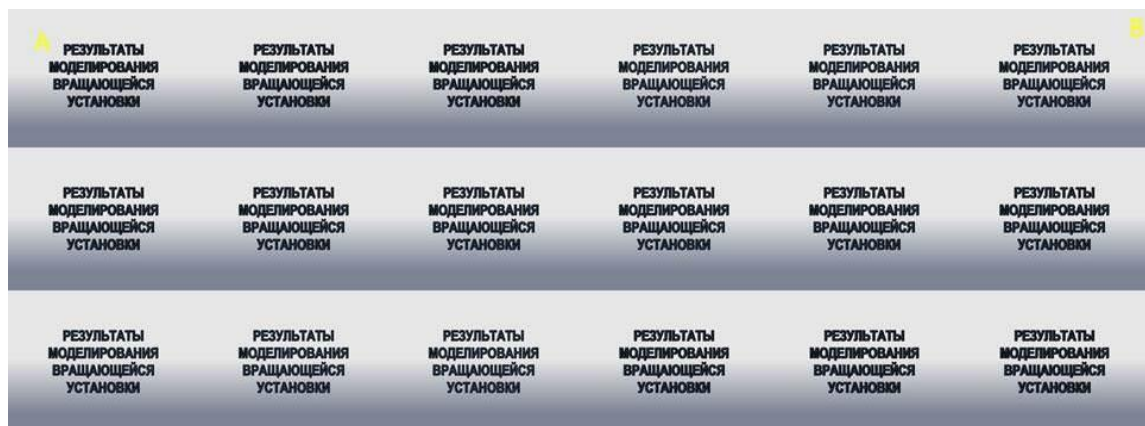

Fig. 7. Variants of gluing frames with a sequence of 1-9 (A) and reverse – 9-1 (B).

During the experiments, different parameters were varied: the font size, the angle of rotation for each frame in the multi-view presentation, the distance between frames for the linear stereo base.

Due to the lack of a theoretical part in this area, the necessary parameters of the linear and angular shift for the multi-view representation had to be sought by the selection method. Optimal parameters were found at which the transitions between stereo pairs became invisible to the viewer, and the inscription itself acquired volume and became convenient for perception. In particular, maximal difference between the first and ninth frames should be no more than 2.4 mm with a step of 0.3 mm, the angle of rotation decreases from 9 degrees to 1, with a step of 1 degree.

It was found that these parameters, in principle, remain the same for very different fonts. A similar effect was obtained for the inscription, where the font size was 32, and the most successful variant of the inscription, where the font size was 66 (Fig. 8 and Fig. 9).

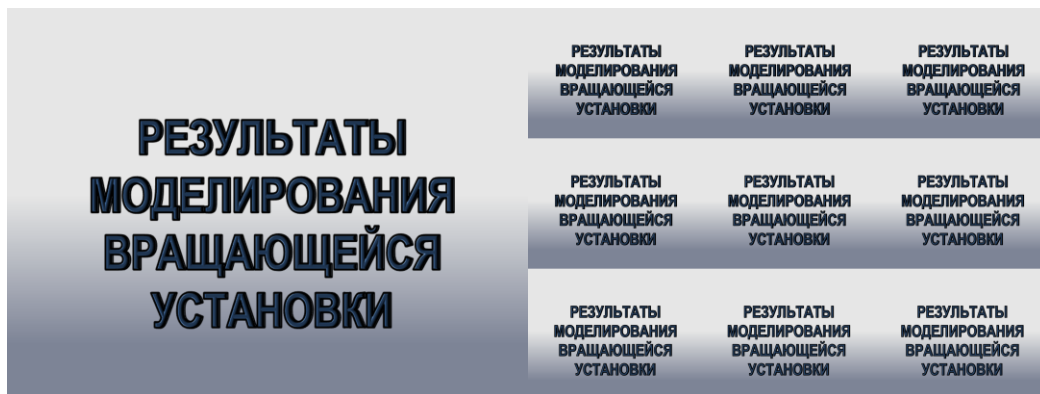**Fig. 8.** Constructing a stereo image of an inscription with a font size of 32.

Fig. 9. Constructing a stereo image of an inscription with a font size of 66.

3. Depth maps

When creating a multi-view video for autostereoscopic monitors, there is a need to store a large amount of data – video streams for each angle. Despite the fact that modern methods of digital video compression can effectively take into account temporal and spatial redundancy, the amount of data with multi-angle video will increase many times. This is especially critical for autostereoscopic monitors.

One of the effective ways to solve the problem of large amounts of data is to use the so-called $2D + Z$ format. Any conventional ($2D$) image can be matched with information about the distance of each pixel from the observer (Z-coordinate). Such a representation of the image is called the "$2D + Z$ format", and the Z coordinate plane is called the "depth map". It can be represented as a monochrome image. In the depth map, grayscale indicates the distance of the image points from the observer. That is, the closest point to the viewer will be white, and the farthest – black.

The idea underlying the construction of a depth map over a stereopair is fairly obvious. For each point on one image, a pair point is searched for it on another image. And by a pair of corresponding points, coordinates of their pre-image in three-dimensional space can be determined. Having the same three-dimensional coordinates of the pre-image, the depth is calculated as the distance to the camera plane.

A pair point must be sought on the epipolar line [13]. Accordingly, to simplify the search, the images are aligned so that all epipolar lines are parallel to the sides of the image (usually horizontal). This image alignment process is called *rectification*.

After the images are rectified, the corresponding pairs of points are searched. The simplest method is as follows: for each pixel of the left image with coordinates (xo, yo), a pixel is searched for in the right image. In this case, it is assumed that the pixel in the right picture should have coordinates ($xo − d$, yo), where d is a quantity called *disparity*. The search for the corresponding pixel is performed by calculating the maximum of the response function,

which can be, for example, the correlation of neighborhoods of pixels. The result is a *disparity map* [14].

This algorithm was used to build depth maps for individual letters and whole labels. To create depth maps, a ready-made software package 3D Stereoscopic Photography [14] was used, created by the author of the package for processing real stereo photographs. This package allows one to create a depth map from the right and left parts of the stereo pair. The interactive part of this complex allows one to set on a stereo pair both the point of the object being visualized closest to the observer and the one furthest from the observer. In this case, both points are indicated both on the left side of the stereo pair and on the right, thereby providing parallax data for further batch processing for the subsequent creation of a depth map.

The algorithm for constructing depth maps is as follows. First, we create a 3D model of the text label, then a stereo pair is constructed from the constructed frame with 3D text using a linear or angular stereo base. Then, using the procedures of the software package [14], the points closest to the observer and farthest from the observer are noted on the stereo pair. This allows us to create depth maps using procedures [14].

To create a depth map, the desired object must have a pronounced depth and volume. If they are insufficient, then it will be impossible to create the necessary image using the algorithm outlined above. For example, a variant of the text that is suitable for a multi-view presentation (Fig. 6) turned out to be completely unsuitable for constructing a depth map due to insufficient letter volume.

Figure 10 shows the variant when one letter was chosen as an experimental sample, enlarged so that the image had a pronounced volume.

Fig. 10. Constructing a single letter depth map.

As a result, it turned out that to construct a depth map, it is enough to have only an angular stereo base (Fig. 11). In contrast to the multi-view representation, the combination of linear and angular stereo-base did not give a positive result when plotting depth maps. It was assumed that the presence of a linear stereo base is not necessary and does not affect the construction efficiency.
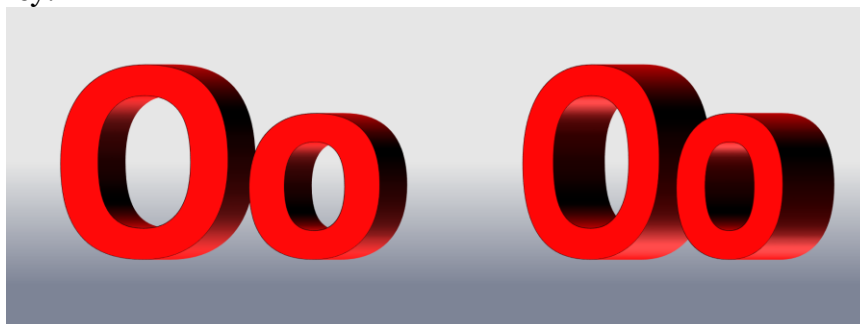

Fig. 11. Constructing of angular stereo base to create a depth map.

As mentioned above, at each of the frames the point farthest from the viewer and the one closest to it is noted. Then the offset distance of the object in the frames is calculated – in this case for two letters – and then a depth map is constructed (Fig. 12).

Fig. 12. Constructing a depth map for left frame of stereo pair.

Depending on the need, it is possible to build a depth map for any of the frames – left or right – then a 3D object will be displayed on the autostereoscopic monitor in the position for which the depth map was built (Fig. 12 and Fig. 13). Moreover, if the right view is used, the depth map is inverse to the depth map for the left view, that is, the closest point to the observer becomes black, and the farthest white.
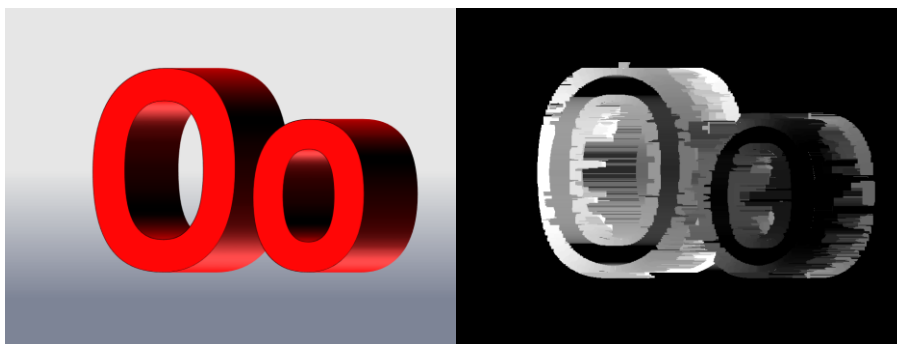

Fig. 13. Constructing a depth map for right frame of stereo pair.

As practice has shown, the construction of a depth map is also possible for the entire inscription, provided that it has a sufficient size and volume (Fig. 14).

Fig. 14. Constructing a depth map for the whole inscription.

As a result of all experiments, a visible stereo effect was achieved and the conditions necessary for constructing depth maps were identified.
However, in the process, the following problem was discovered – the presence of artifacts clearly visible on the autostereoscopic monitor. In the above images, they are also visible, only on depth maps.

The reasons are that the algorithms for smoothing the constructed depth map used by the author in the software package [14] used in this study work well mainly with real stereo photographs when the visualization object does not contain sharp edges with a contrasting color change. In our case, the object of visualization is text, most often possessing such properties. Thus, further studies suggest the selection and application of the most optimal smoothing algorithms.

4. Conclusion

The results of constructing textual information for an autostereoscopic monitor are presented using two methods for constructing a stereo image – a multi-view representation and depth maps. In practical experiments, optimal parameters for constructing stereo images of the text with the most visible stereo effect for both methods are empirically obtained.

The purpose of this research is to develop an effective technology for constructing stereo images of the text to create stereo presentations. In the longer term, we suggest improving the quality of the created depth maps – smoothing and eliminating artifacts – as well as solving the problem of combining images and text information in one stereo frame. The aim of the study as a whole is to create a technology that allows one to quickly and efficiently create full-fledged animated stereo presentations of scientific results for various types of stereo devices.

5. References

[1] Generation of Stereo-Presentations in Photorealistic Rendering and Scientific Visualization /Andreev S.V.[et al] // Keldysh Institute preprints, 2010. № 61. 16 p. http://library.keldysh.ru/preprint.asp?id=2010-61[in Russian].

[2] Synthesis of photorealistic three-dimensional images in modern presentation systems / Andreev S.V.[et al] // Software&Systems 2007. № 3. p. 37-40. [in Russian].

[3] Andreev S., Filina A. Using stereo presentations for visualization of scientific calculations results / Scientific Visualization. 2012. v.4. № 1. p.12-21.

[4] Hardware-software complex of 3D presentations based on a virtual studio and virtual environment / Vandanov V.G. [et al] // Proceedings of the 1-st international conference «3D visualization of scientific, technical and social reality. Cluster technologies of modeling» Izhevsk. 2009. P.73-77. [in Russian].

[5] Mezhenin A.V., Tozik V.G. 3D Visualization using the stereo image effect // Proceedings of the 2-nd international conference «3D visualization of scientific, technical and social reality. Cluster technologies of modeling» Izhevsk. 2010. [in Russian].

[6] Mikhaylyuk M.V., Huraskin I.A. Synthesis of stereo images for virtual reality systems using an optical tracking system / Software&Systems 2006. № 3. p. 10-14. [in Russian].

[7] Andreev S.V. [et al] / Modelling and visualization of blade assembly with complicated shape for power turbine // Scientific Visualization. 2015. v.7. № 4. p.1-12

[8] Torgashev M.A., P.Y. Timokhin. The technology of stereo video files' synthesis for the system of 3D real-time visualization. Software Products and Systems, 2012, № 3, pp. 74-80. [In Russian]

[9] Mikhaylyuk M.V., Maltsev A.V. Timokhin P.Yu. The methods of 3D stereo visualizationof data obtainedin simulation of unstable oil displacement from porous media / Proceedings of Scientific Research Institute for System Analysis RAS, 2018, v.8, N 2, p. 125-129. [in Russian].

[10] Visibility Experiment and Evaluation of 3D Character Representation on Mobile Displays / Hiromu Ishio [et al.] // C. Stephanidis (Ed.): Posters, Part II, HCII 2011, CCIS 174, pp. 46–51, 2011.

[11] Andreev S.V., Bondareva N.A. Constructing a representation of textual information in stereo presentations // Proceedings of the 28-th International Conference GraphiCon-2018, Tomsk, Russia, 23-28 September 2018, p. 86-89. [in Russian].

[12] S.V.Andreev, A.E.Bondarev, V.A.Galaktionov, N.A.Bondareva (2018) The problems of stereo animations construction on modern stereo devices. Scientific Visualization 10.4: 40 - 52, DOI: 10.26583/sv.10.4.04

[13] Kalinichenko A.V., Sveshnikova N.V., Yurin D.V. Epipolar geometry and evaluation of its reliability from the results of the restoration of the three-dimensional scene by factorization algorithms // Proceedings of the 16-th International Conference GraphiCon-2006, 2006. [in Russian].

[14] 3D Stereoscopic Photography, http://3dstereophoto.blogspot.com/ last access September 01 2019.