Computational Mathematics and Modeling, Vol. 23, No. 2, April, 2012

IMAGE SYNTHESIS FOR AUTOSTEREOSCOPIC SYSTEMS

A. A. Luk'yanitsa

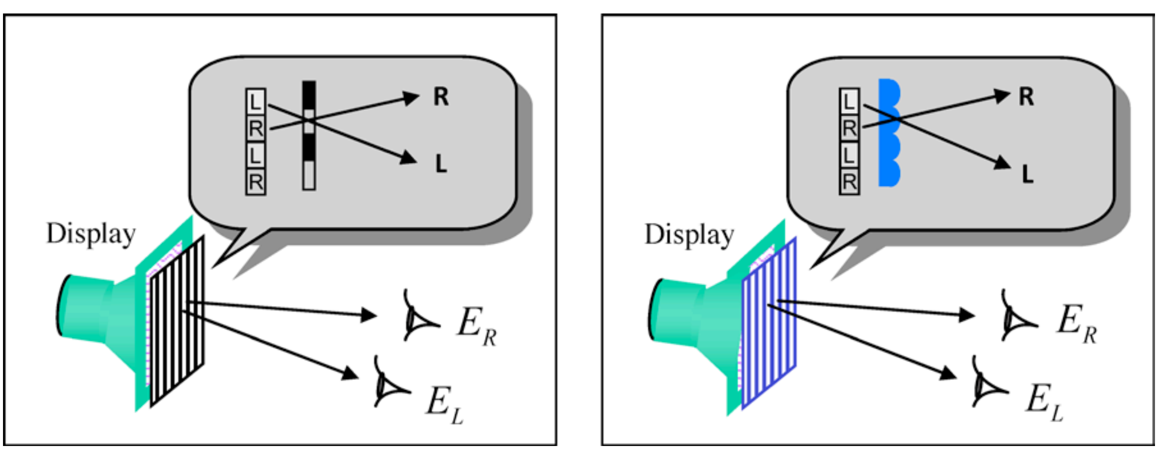
UDC 517.988.38:535.317.2

The article considers image synthesis for autostereoscopic systems — systems that enable the viewer to perceive three-dimensional images without any special devices, such as glasses. The general image synthesis problem for such systems is stated and an efficient algorithm is proposed for its solution. Some applications of the proposed methodology for two types of units are described: a stereo display and a stereo projection system. We examine the optimization of technical system parameters with the objective of enlarging the region where the viewer sees a three-dimensional image.

Keywords: autostereoscopic system, image synthesis, functional minimization.

Introduction

The present article considers image synthesis for autostereoscopic systems, i.e., systems that enable the viewer to perceive three-dimensional images without any special devices, such as glasses (also known as naked-eyes systems). To see a three-dimensional perspective, the viewer's left and right eyes should be trained at different images, which correspond to different viewing points of the scene. This is easiest to achieve by adding a parallax barrier, i.e., place a special mask between the screen and the viewer, which prevents the image intended for the left eye from reaching the right eye, and vice versa. The simplest mask is a transparent film covered with opaque vertical bands — Fig. 1a. Masks constructed from miniature cylindrical lenses (lenticular lenses) are also used (Fig. 1b). Given the mechanisms of stereoscopic vision, these devices have a periodic structure – the images for the left and the right eye are separated by the bands. The shortcomings of these devices include reduction of image brightness, appearance of "false" zones where the left eye sees the image intended for the right eye and vice versa, and very narrow stereo-effect zones.



(a) **(b)**

Fig. 1. Autostereoscopic systems with parallax barrier: (a) a slit mask, (b) a lenticular raster.

Faculty of Computational Mathematics and Cybernetics, Moscow State University, Moscow, Russia; e-mail: andrei.luk@gmail.com.

Translated from Prikladnaya Matematika i Informatika, No. 37, pp. 80–97, 2011.

1046–283X/12/2302–0195

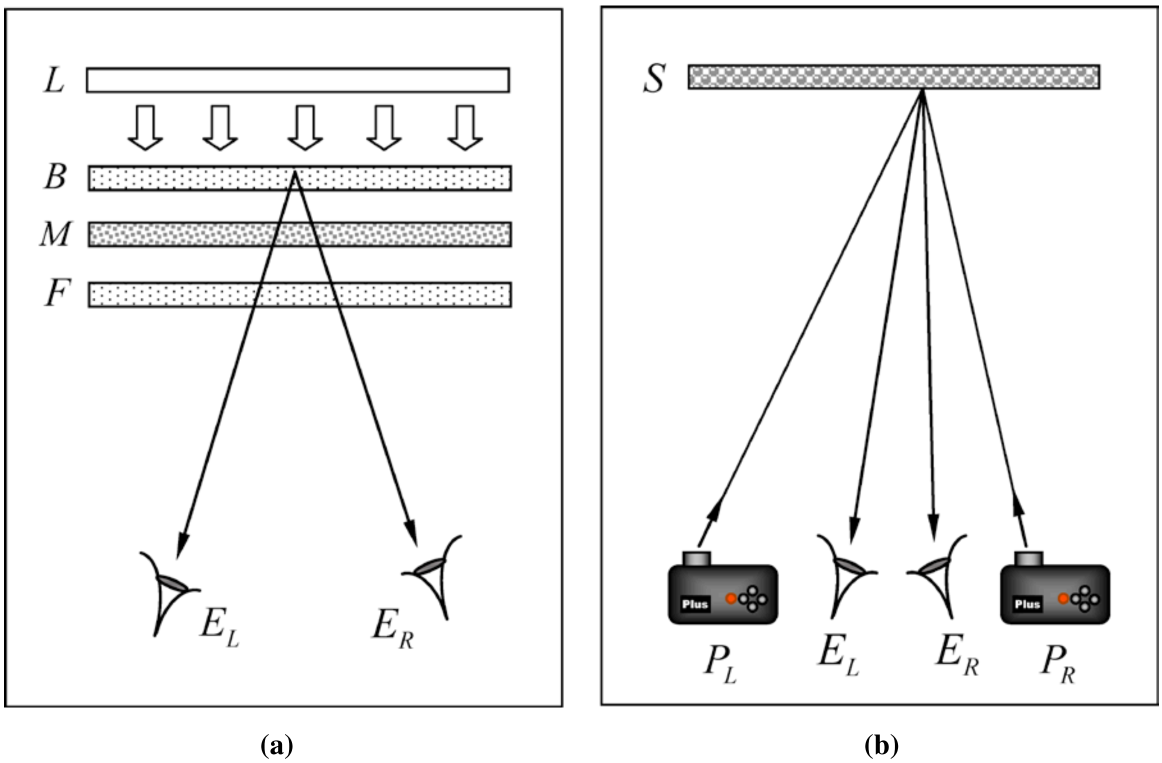


Fig. 2. Systems that generate continuous 3D images: (a) stereo display, (b) projection system.

In the present article we consider a stereo display and a projection system in which the 3D effect is achieved by integrating several specially prepared continuous images and using light scattering. This technique produces a stereo image in a fairly wide spatial region. A stereo display [1, 2] consists of two liquid crystal (LC) panels F and B with a diffuser M between them.(Fig. 2a). Specially computed images are displayed on the two panels, acting as masks for the light that reaches the viewer's eyes E_L and E_R from the backlight L. Since the LC panels are separated by a certain distance in depth, the images reaching the left and the right eye pass through different mask pixels and are therefore distinct. The diffuser has a dual function. First, it eliminates the moiré effect produced by the periodic structure of the LC panels — the pixels are separated by opaque black bands. Second, the diffuser makes it possible to increase the viewing angle. When the viewer's position shifts, other mask pixels are superimposed and the image changes. The diffuser slightly "fuzzes" the rear mask, so that small deviations of the viewer from the chosen position do not disrupt stereo perception.

A projection system (3) consists of two (or more) standard projectors and a retro-reflective screen (Fig. 2b), which reflects light strictly in the direction of the source. However, many retro-reflective materials have diffusive properties and therefore reflect light into a certain solid angle around the direction of the source, with exponential decay at the periphery. This property may be exploited to create a 3D light field in a fairly wide zone. The scattering law for retro-reflective screens is adequately modeled by the normal distribution

$$K(\alpha) = P \cdot \exp\left(-\frac{\alpha^2}{\sigma^2}\right), \tag{1}$$

where P is the amount of light absorbed by the screen, α is the angle between the light source and the viewer, and σ is the parameter characterizing the scattering properties of the screen. Since the viewer's left and right eyes see images reflected from different projectors at different angles, it is possible to select projector images so that the viewer perceives the desired 3D scene.

These stereo systems create different images at different viewing angles and the reproduction of a desired 3D scene therefore requires computing (synthesizing) the images that should be applied to the system input. Such images are not determined uniquely, and we have to select the images that ensure the widest region of stereo vision. The present article aims to solve this problem. It is organized as follows. In Section 1 we present a general statement of the image synthesis problem for autostereoscopic systems. In Section 2 we describe an efficient algorithm for solving this problem. Sections 3 and 4 present applications of the proposed image synthesis methodology for a stereo display and a stereo projection system. The last section, Section 5, examines the choice of optimal technical parameters for these systems with the objective of enlarging the region where the viewer sees a 3D image.

1. Statement of the Problem

Given is an optical stereo system S for reproduction of images whose design parameters are known. For the stereo display the known parameters include the distance between the panels, screen size, and diffuser properties. We refer to these parameters as the system internal parameters and denote their entire set by H. Take a Cartesian coordinate system and place the stereo system S at a given point in space $\vec{X}_S$. To allow the viewer to see a 3D picture, the stereo system must have the following property. If $n \geq 2$ images $I_1, \ldots, I_n$ computed in a special way are applied to the system input, then the viewer's eyes positioned at points with the coordinates $\vec{X}_L$ and $\vec{X}_R$ should see different images

$$L = S\left(\vec{X}_L, \vec{X}_S, I_1, \ldots, I_n, H\right), \qquad R = S\left(\vec{X}_R, \vec{X}_S, I_1, \ldots, I_n, H\right). \tag{2}$$

For the viewer to see the desired three-dimensional picture, these images should match a given stereo pair $\left(I_L, I_R\right)$, i.e., the two images that are projections of the original scene from two viewing points (these two images are usually taken with a stereo camera). Since the position of the viewer relative to the stereo system and the system internal parameters are fixed, the desired effect can be achieved by altering only the images $I_1, \ldots, I_n$, which are determined by minimizing the discrepancy functional

$$\Phi_1\left(I_1, \ldots, I_n\right) = \frac{1}{2}\left(L\left(I_1, \ldots, I_n\right) - I_L\right)^2 + \frac{1}{2}\left(R\left(I_1, \ldots, I_n\right) - I_R\right)^2 . \tag{3}$$

Assume that all the images are discrete, of size N_w and N_h in the horizontal and the vertical direction respectively. Then the functional (3) may be written as

$$\Phi_1\left(I_1, \ldots, I_n\right) = \frac{1}{2}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\left(\left(L(i,j) - I_L(i,j)\right)^2 + \left(R(i,j) - I_R(i,j)\right)^2\right). \tag{4}$$

Since the intensity of the images to be rendered on the display or the projector is bounded by the range $[0, 255]$, the sought images should satisfy the constraints

$$0 \leq I_k(i,j) \leq 255 , \qquad k = 1, \ldots, n, \quad i = 1, \ldots, N_h, \quad j = 1, \ldots, N_w . \tag{5}$$

In principle, we can accept any set of images $I_1, \dots, I_n$ satisfying the constraints (5) that minimize the functional Φ_1. However, our problem requires choosing the smoothest of all admissible solutions, because this will enable the viewer to perceive a 3D picture not only at the computed point, but also in some neighborhood around it. To find the smoothest solution, we have to add another functional to the discrepancy functional:

$$\Phi_2(I_1, \dots, I_n) = \frac{1}{2}\sum_{k=1}^{n}\sum_{i=1}^{N_h-1}\sum_{j=1}^{N_w-1}\left((I_k(i+1,j)-I_k(i,j))^2+(I_k(i,j+1)-I_k(i,j))^2\right). \tag{6}$$

This functional is moreover a regularizing increment for the functional Φ_1.

The image synthesis problem thus involves minimizing the functional

$$\Phi(I_1, \dots, I_n) = \Phi_1(I_1, \dots, I_n) + \alpha\Phi_2(I_1, \dots, I_n) \tag{7}$$

subject to the constraints (5).

2. Image Synthesis Algorithm

Problem (7), (5) is a constrained minimization problem for a quadratic functional with inequality constraints. To solve the problem, we can construct a Lagrange functional and find the minimum value from the Kuhn–Tucker conditions [4]. However, our problem is reducible to an unconstrained minimization problem. To this end we make the change of variables

$$I_k = \frac{255}{1+\exp(-J_k)}. \tag{8}$$

With this change of variables, I_k lies in the interval $[0, 255]$, i.e., conditions (5) are satisfied automatically, and the new variables J_k may vary over the entire real axis. The functional (7) can be efficiently minimized by gradient descent with an inertial term:

$$J_k(t+1) = J_k(t) - \eta\frac{\partial\Phi(t)}{\partial J_k} - \mu\frac{\partial\Phi(t-1)}{\partial J_k}. \tag{9}$$

Here η and μ are the iteration parameters characterizing the speed of the gradient descent process, t is a fictitious parameter denoting the iteration number. The inertial term allows for the gradient change from the preceding step, thus avoiding getting trapped in a local minimum. The image quality is characterized by the percentage difference between the generated images (L, R) and the specific stereo pair (I_L, I_R):

$$E = \frac{100\%}{255}\sqrt{\frac{1}{2N_hN_w}\sum_{i=1}^{N_h}\sum_{j=1}^{N_w}(L(i,j)-I_L(i,j))^2+\frac{1}{2N_hN_w}\sum_{i=1}^{N_h}\sum_{j=1}^{N_w}(R(i,j)-I_R(i,j))^2}. \tag{10}$$

E is called the image synthesis error.

Since all digital imaging devices, such as displays or projectors, have essentially discrete structure, a stereo system is conveniently represented as a neural network N where neurons correspond to image pixels and the connections W between the neurons are fixed; they model the propagation of light rays in accordance with the set of internal parameters H and the position of the viewer $\vec{X}$ relative to the position of the system $\vec{X}_S$, $W = W\left(\vec{X},\vec{X}_S,H\right)$ [5, 6]. The images $I_1,\ldots,I_n$ are applied to the neural network input, and once the signal has traversed the network, the images formed at the output are displayed to the viewer:

$$L = N\left(W\left(\vec{X}_L,\vec{X}_S,H\right),I_1,\ldots,I_n\right), \qquad R = N\left(W\left(\vec{X}_R,\vec{X}_S,H\right),I_1,\ldots,I_n\right). \tag{11}$$

For a neural network, the functional (7) is an error function, the change of variables (8) corresponds to the introduction of a nonlinear transfer function for the neurons, and the gradient descent process (9) is a network tuning procedure called backward error propagation [7].

3. Image Synthesis for Stereo Display

Take an orthogonal coordinate system attached to the viewer. Point the Ox axis to the right, the Oy axis down, and the Oz axis so as to complete a right-handed triad. Assume that the origin is at the point half-way between the viewer's eyes. In this coordinate system, the left eye is at the point $\left(x_L, 0, 0\right)$ and the right eye at the point $\left(x_R, 0, 0\right)$, with $x_L = -x_R$. Assume that the screens are perpendicular to the Oz axis and it passes through their centers; the front screen is at a distance D from the origin, and the inter-screen distance is d (Fig. 3).

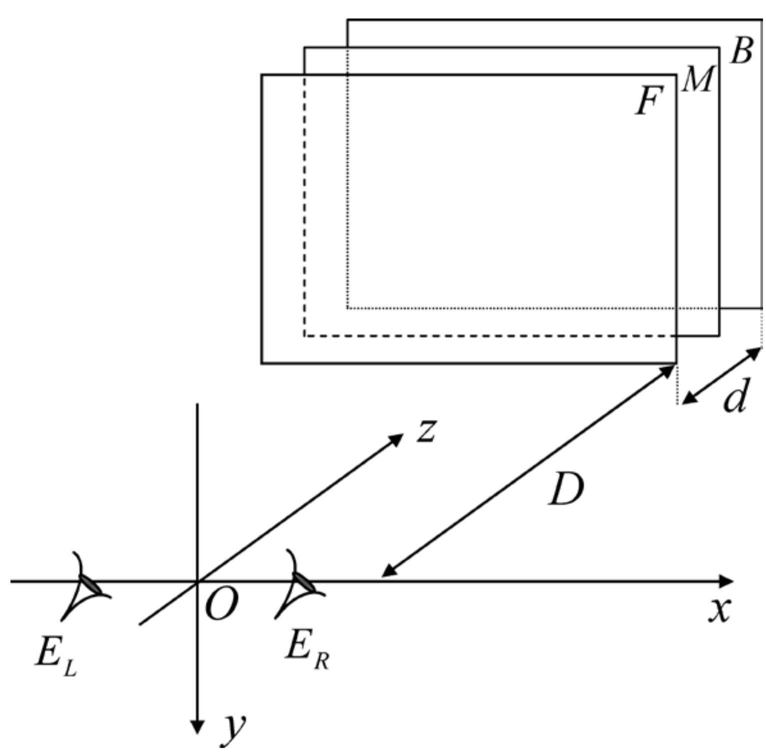


Fig. 3. Formation of stereo image by stereo display. Here E_L and E_R are the viewer's left and right eye, F and B are the front and the back LC panel, M is the diffuser, D is the distance from the observer to the near panel, d is the inter-panel gap.

The light from the backlight passes through the rear screen, undergoes slight diffusion in the diffuser, and then passes through the front screen. In the end, the left and the right eyes see the following images:

$$L(x,y) = I_1(x,y)\iint_{\Omega} K(\xi - ax + bx_L, \eta - ay) I_2(\xi,\eta)\, d\xi d\eta, \tag{12}$$

$$R(x,y) = I_1(x,y)\iint_{\Omega} K(\xi - ax + bx_R, \eta - ay) I_2(\xi,\eta)\, d\xi d\eta.$$

Here $I_1(x,y)$ and $I_2(x,y)$ respectively are the image intensities on the front and the back screen, $a = (D+d)/D$, $b = d/D$, and the kernel modeling the diffuser effect has the form

$$K(t_1,t_2) = p \cdot e^{-\frac{(t_1^2+t_2^2)}{\sigma^2}} \tag{13}$$

Here p is the absorption coefficient and σ a parameter characterizing the diffuser scattering properties. Integration in (12) is over the region Ω that the image occupies on the screen.

The problem has been solved by the method described in Section 2. Figure 4 plots the variation of the image synthesis error E as a function of the iteration number t for various values of the parameter η. For $\eta = 0.001$ the convergence is slow, and for $\eta = 0.005$ the error fluctuates around 25 %.

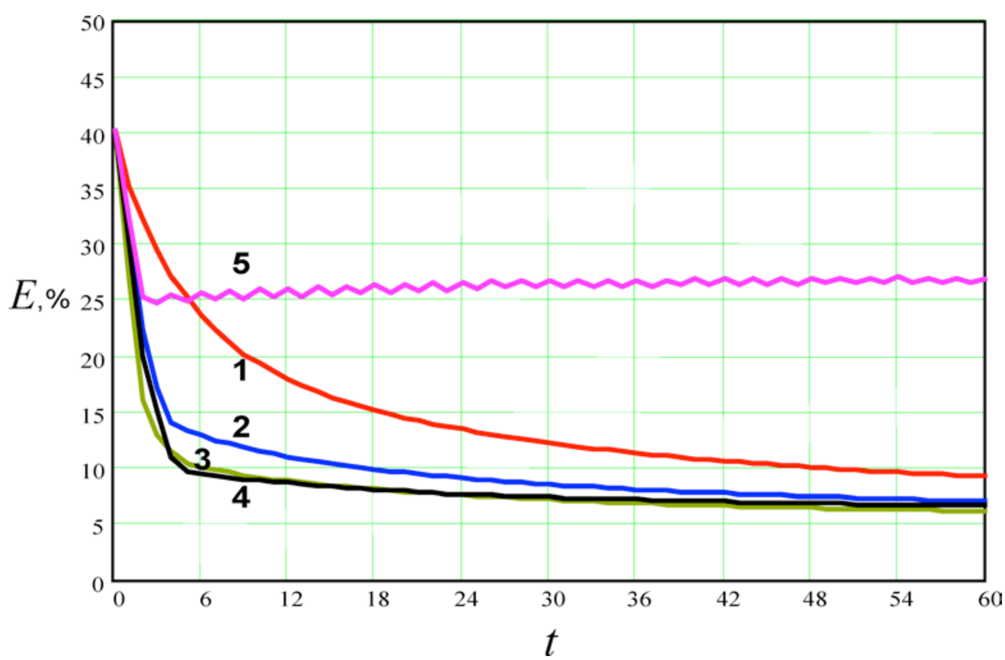


Fig. 4. Image synthesis error E versus the iteration number for various values of the parameter η: (1) $\eta = 0.001$, (2) $\eta = 0.002$, (3) $\eta = 0.003$, (4) $\eta = 0.004$, (5) $\eta = 0.005$.

The optimal value is $\eta = 0.003$. For the parameter μ characterizing the contribution of the inertial term we experimentally obtained an optimal value equal to η.

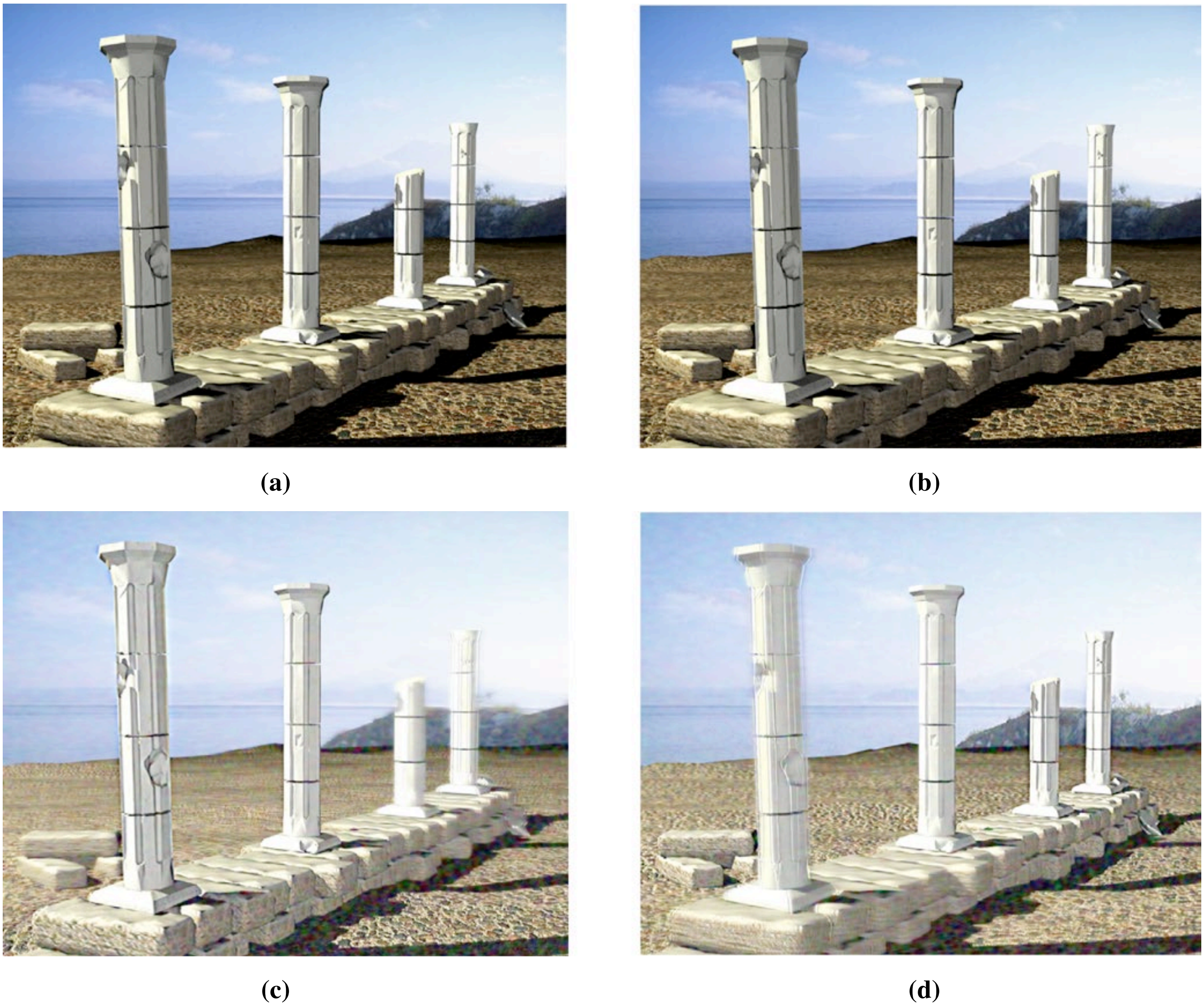


Fig. 5. Results of a numerical experiment: (a), (b) initial stereo pair; (c), (d) computed images I_1 and I_2 to be delivered to the front and the back screen, respectively.

Figure 5 presents the simulation results. The top row (a, b) shows the original stereo pair (I_L, I_R), and the bottom row (c, d) shows the computed images I_1 and I_2 that are delivered to the front and the back screen, respectively.

4. Image Synthesis for Projection System

Assume that the reflective screen is placed vertically. Introduce a Cartesian coordinate system with its origin at the center of the screen. The horizontal axis Ox and the vertical axis Oy are in the screen plane. The axis Oz is directed toward the projectors, as shown in Fig. 6.

Let the viewer's eyes be at the points with the coordinates $\vec{E}_L$ and $\vec{E}_R$, and the projectors at the points $\vec{P}_1$ and $\vec{P}_2$, receiving the images I_1 и I_2. Then the viewer sees with the left and the right eye the images L

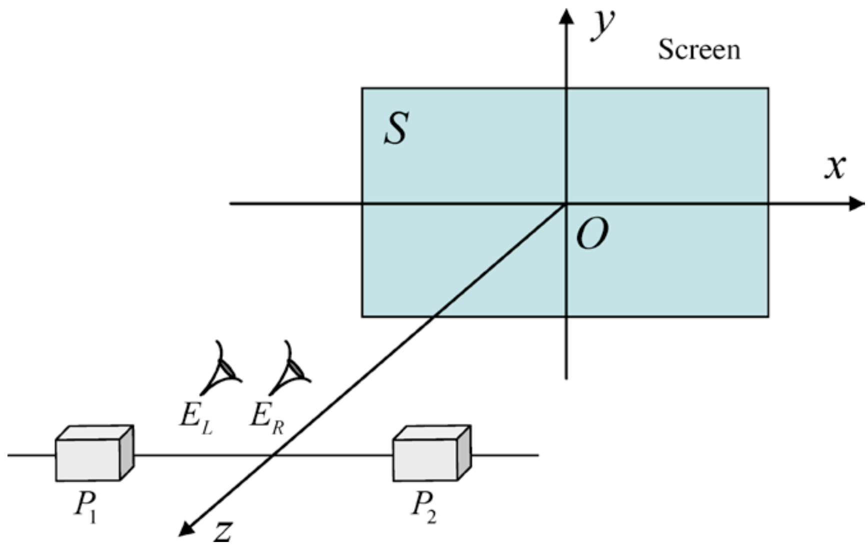


Fig. 6. Projection system: S is the screen, P_1 and P_2 are the projectors, E_L and E_R are the viewer's eyes.

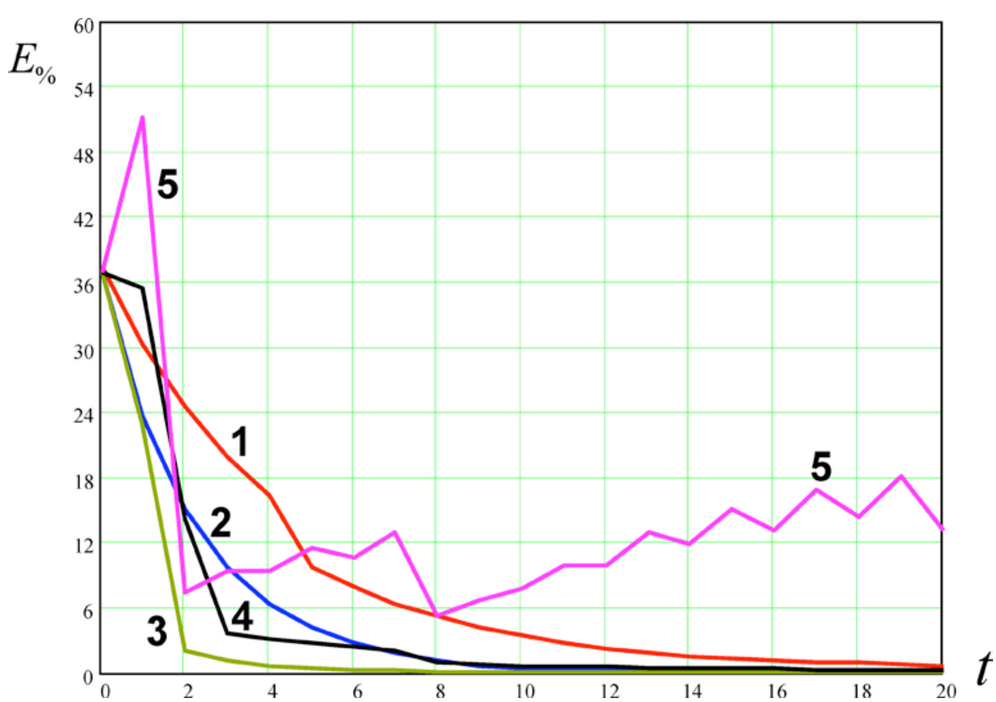


Fig. 7. Image synthesis error E as a function of the iteration number t for various values of the parameter η: (1) $\eta = 0.05$, (2) $\eta = 0.1$, (3) $\eta = 0.9$, (4) $\eta = 1.0$, (5) $\eta = 1.1$.

and R, which are respectively computed by the following formulas:

$$L(x,y) = I_1(x,y) \cdot K\left(\alpha_{11}\left(x,y,\vec{E}_L,\vec{P}_1\right)\right) + I_2(x,y) K\left(\alpha_{12}\left(x,y,\vec{E}_L,\vec{P}_2\right)\right), \quad (14)$$

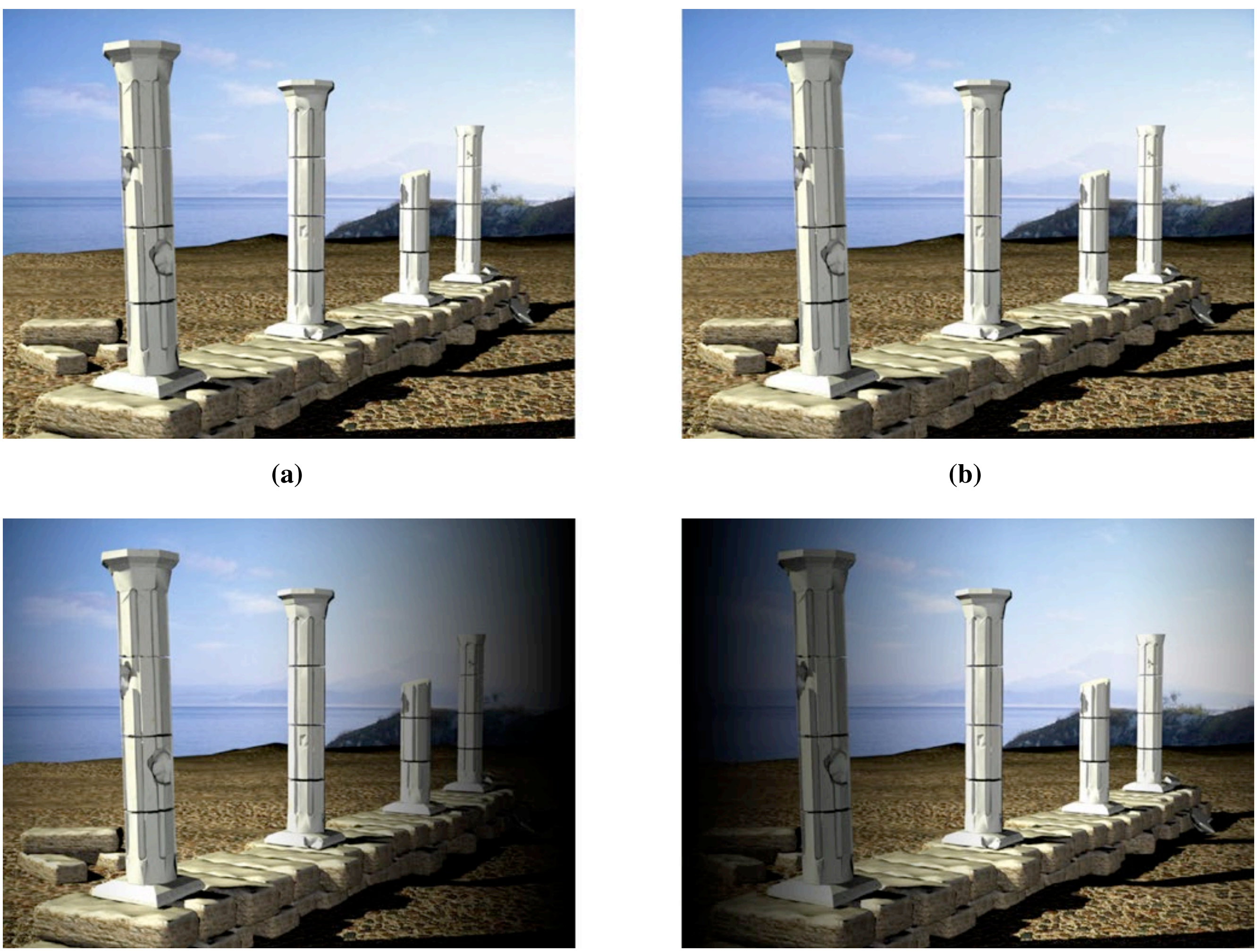
$$R(x,y) = I_1(x,y) \cdot K\left(\alpha_{21}\left(x,y,\vec{E}_R,\vec{P}_1\right)\right) + I_2(x,y) K\left(\alpha_{22}\left(x,y,\vec{E}_R,\vec{P}_2\right)\right),$$

where α_{ij} is the angle between the incident light ray and the reflected ray, i.e., between the ray that leaves the projector j, is reflected from the screen at the point (x, y), and reaches the viewer's eye i; $K(\alpha)$ is the reflection law for the retro-reflective screen,

$$K(\alpha) = P \cdot \exp\left(-\frac{\alpha^2}{\sigma^2}\right), \tag{15}$$

where P is the absorption of light by the screen and σ a parameter characterizing the scattering properties of the screen.

The method of Section 2 has been applied to synthesize the projector images from the given stereo pair (I_L, I_R). Figure 7 plots the variation of the image synthesis error E as a function of the iteration number t for various values of the parameter η. For $\eta = 0.05$ the convergence is slow, and starting with $\eta = 1.1$ the process does not converge at all.



(a) (b) (c) (d)

Fig. 8. Results of numerical computations: (a), (b) the original stereo pair (I_L, I_R); (c), (d) the computed images I_1 and I_2 to be delivered to the left and right projectors.

The optimal value of this parameter is $\eta = 0.9$, when the process converges in 4–5 iterations, achieving an error of less than 1 %. The parameter μ characterizing the contribution of the inertial term was taken equal to η.

Figure 8 shows the simulation results. In the top row (a, b), we have the original stereo pair (I_L, I_R), and in the bottom row (c, d) the computed images I_1 and I_2 that should be applied to the projectors.

5. Optimization of System Internal Parameters

Image reproduction quality in stereo systems depends on the set of technical parameters H adopted in manufacturing. For stereo display the main parameters are the pixel aperture, the inter-panel distance, and the diffuser scattering characteristic. For a projection system important parameters include the screen scattering parameter, the separation between the projectors, and the distance of the projectors from the screen.

We introduce a Cartesian coordinate system with the center at the viewer's position, the axis Ox pointing to the right, the vertical axis Oy directed downward, and the axis Oz toward the screen (Fig. 9).

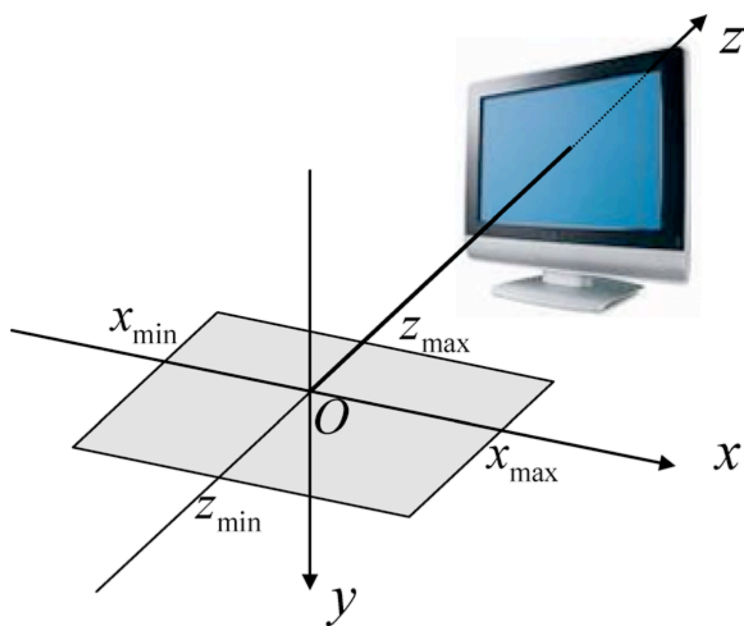


Fig. 9. When viewing the stereo system screen, the viewer is free to move within the region $[x_{\min}, x_{\max}] \times [z_{\min}, z_{\max}]$.

Given is the stereo pair (I_L, I_R). Fix the values of the parameters H and compute the image (L, R). Assume that the viewer shifts to some point with the coordinates (x, z). The quality of the stereo image at this point is characterized by the following measure:

$$e_{xz} = \sqrt{\frac{1}{2N_h N_w} \sum_{i=1}^{N_h} \sum_{j=1}^{N_w} \left(L(i,j) - I_L(i,j)\right)^2 + \frac{1}{2N_h N_w} \sum_{i=1}^{N_h} \sum_{j=1}^{N_w} \left(R(i,j) - I_R(i,j)\right)^2} \cdot \frac{100\%}{255}, \qquad (16)$$

where N_h and N_w is the number of pixels in the image along the horizontal and the vertical respectively. In effect, e_{xz} is equal to the stereo image reproduction error, expressed in percent. The smaller this error, the more accurate is the reproduction of the image by the system for the given position of the viewer. It is important to ensure that, as the viewer shifts within certain limits, he continues to see a stereo image close to the desired one. Therefore our task is to choose the system parameters H that maximize the viewing region $P(H)$ where the desired stereo image is properly seen. To compute $P(H)$, we superimpose a uniform grid on the region of the viewer's possible positions $[x_{\min}, x_{\max}]\times[z_{\min}, z_{\max}]$:

$$x_j = x_{\min} + (j-1)\frac{(x_{\max} - x_{\min})}{N_x - 1}, \quad j = 1, \dots, N_x, \tag{17}$$

$$z_i = z_{\min} + (i-1)\frac{(z_{\max} - z_{\min})}{N_z - 1}, \quad i = 1, \dots, N_z,$$

where N_x and N_z is the number of grid nodes along x and z, respectively.

Real-life tests have shown that viewers do not notice defects in the stereo image if the error (16) does not exceed some $e_{\min}$. We therefore numerically characterize the size of the stereo zone by the ratio of the area of the region where the error e_{xz} does not exceed the specified limit $e_{\min}$ to the total viewing area $[x_{\min},x_{\max}]\times[z_{\min},z_{\max}]$. In case of discrete moves of the viewer over the grid (17), this ratio is computed in the following way. Define the grid function

$$p_{xz} = \begin{cases} 1 & \text{if} \quad e_{xz} \le e_{\min}, \\ 0 & \text{if} \quad e_{xz} > e_{\min}, \end{cases} \tag{18}$$

then

$$P(H) = \frac{1}{N_x N_z}\sum_{i=1}^{N_z}\sum_{j=1}^{N_x} p_{ji} \cdot 100\%. \tag{19}$$

We can now introduce a grid over the region of possible variation of the parameters H and compute $P(H)$ at the grid nodes. If the grid spacing is sufficiently small, we can find the optimal parameter values by direct enumeration, or alternatively use the values computed on a coarse grid to construct approximating dependences and thence find the optimal parameter sets.

Figure 10 shows the level lines of the function $f(x,z) = e_{xz}$ in the standard viewing zone for a 17" display with distances of 2 mm and 6 mm between the panels (Figs. 10a and 10b, respectively).

As the distance between the panels increases, the error with the viewer positioned at $(0, 0)$ decreases, while at the same time the error increases at the periphery of the region. The quality of the stereo image reproduced by the display is influenced by the following parameters:

a — the aperture of the display pixels,

d — the distance between the back and the front panels,

σ — the diffuser scattering parameter.

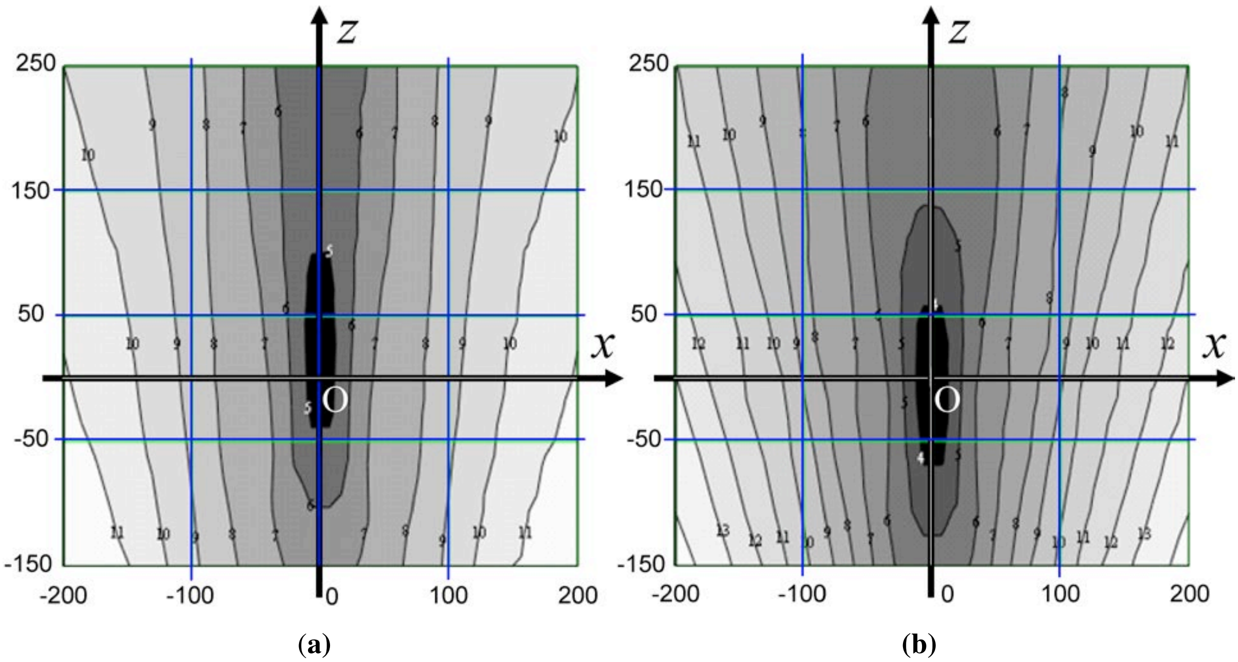


Fig. 10. Level lines of the function $f(x,z) = e_{xz}$ with 2 mm (a) and 6 mm (b) between the panels. The distances along x and z are in millimeters.

Numerical calculations have produced the following approximation formula for the function $P(H)$ characterizing the width of the stereo zone:

$$P(a,d,\sigma) = 36.4 \cdot a^{0.36} e^{-\left(\frac{d-5.01}{4.55}\right)^2} e^{-\left(\frac{\sigma-0.028}{0.19}\right)^2} . \tag{20}$$

This formula has been obtained with $e_{\min} = 5\%$ and it achieves $94\,\%$ accuracy on the full set of computation data. The formula implies that the effect of the aperture on the width of the stereo zone is fairly small; the optimal inter-panel distance is $d \approx 5$ mm, and the optimal scattering parameter is $\sigma \approx 0.03$.

Similar calculations have been carried out for a projection system. Figure 11 shows the level lines of the function $f(x,z) = e_{xz}$ for the following stereo system parameters: screen-to-projector distance $d_S = 4$ m, separation between the projectors $d_P = 20$ cm, screen-to-viewer distance $d_E = 4$ m. The left panel corresponds to $\sigma = 0.03$, the right panel $\sigma = 0.05$.

The area of the stereo zone increases with the increase of the scattering coefficient σ. The technical parameters of the projection system are determined by three main parameters: σ characterizes the scattering properties of the screen material; the other two parameters — the screen-to-projector distance d_S and the projector separation d_P — determine the geometrical characteristics of the projection system. Since the screen properties are given and the screen-to-projector distance is determined by projection theater architecture, the only parameter that allows relatively easy adjustment is the projector separation d_P. It is therefore reasonable to state the parameter optimization problem in the following form: given the parameters σ and d_S, find d_P that maximizes the area of the stereo zone $P(\sigma, d_S, d_P)$. Numerous calculations carried out for various combinations of the system parameters have led to approximate formulas for the projector separation d_P that maximizes

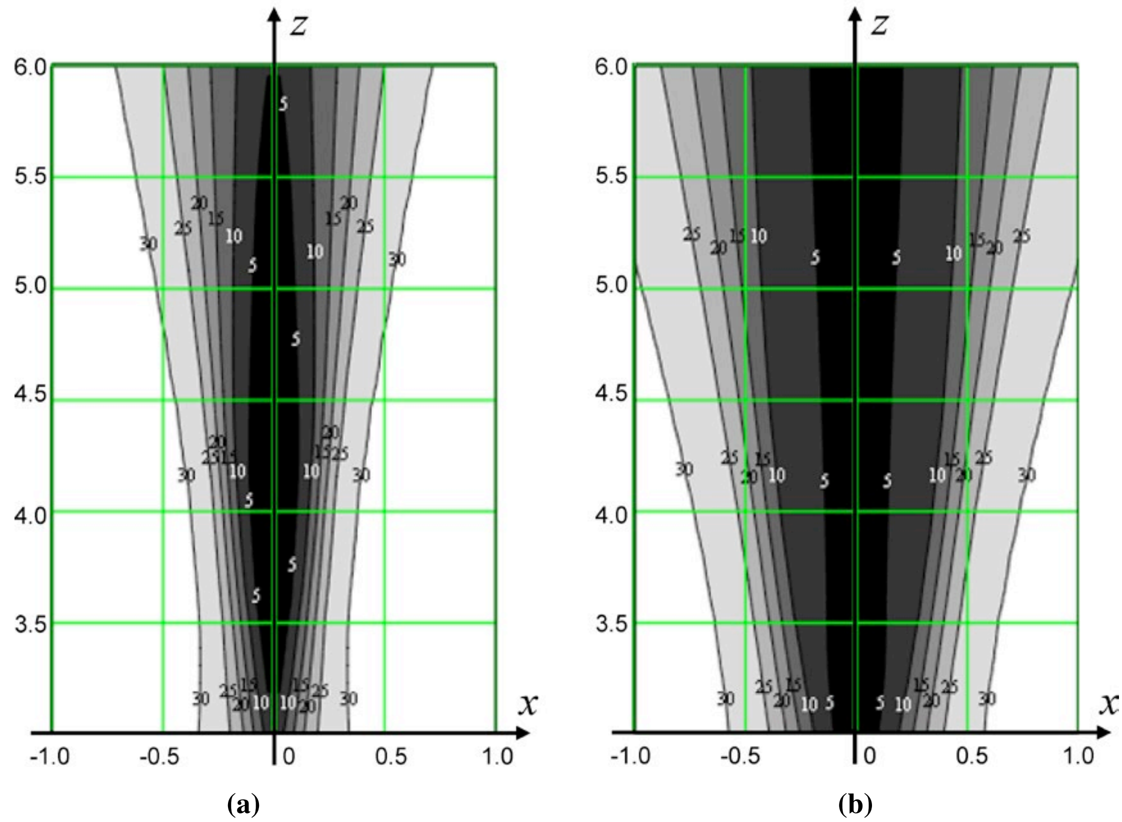


Fig. 11. Level lines of the function $f(x, z) = e_{xz}$ for the following parameter values: d_S = 4 m, d_P = 20 cm, d_E = 4 m, σ = 0.03 (a) and σ = 0.05 (b). Distances in meters.

the size of the stereo zone:

$$d_P \approx (0.6 + 130\sigma)d_S + 120\sigma . \quad (21)$$

Here the screen-to-projector distance d_S is in meters and the projector separation d_P in centimeters. This formula has been obtained with $e_{\min}$ = 5 % and it achieves 97.1 % accuracy on the full set of computation data.

REFERENCES

1. A. Luk'yanitsa and A. Putilin, "A technique for the reproduction of the image of an object," *Russian Agency of Patents and Trademarks*, **RU 2158949 C1** (1999).
2. A. A. Lukyanitsa and A. N. Putilin, "Visualization of three dimensional images and multi aspect imaging," United States Patent 6985290 B2 (2006).
3. A. A. Lukyanitsa and A. N. Putilin, "Three-dimensional image projection employing retro-reflective screens," United States Patent 6843564 B2 (2005).
4. F. P. Vasil'ev, *Numerical Methods for Extremal Problems* [in Russian], Nauka, Moscow (1998).
5. A. A. Luk'yanitsa, "Mathematical modeling of autostereoscopic systems," *VANT*, *Ser. Matem. Modelirovanie Fiz. Protsessov*, No. 1, 74–82 (2007).
6. A. Lukyanitsa, K. Kanashin, and A. Putilin, "Stereodisplay with neural network image processing," *Proc. SPIE Conference*, 20–15 Jan 2002, San Jose, CA.
7. S. Khaikin, *Neural Networks: A Complete Course* [in Russian], Izd. "Vil'yams", Moscow (1996).