S 7-1. *Invited Paper:* 3D Displays: Fundamental Physical Classification for Clarifying Inherent Technical Features

Vasily A. Ezhov
Prokhorov Institute of General Physics, 38 Vavilova str., Moscow 119991, Russia, email: ezhov@3dstereo.ru

Abstract: Four fundamental physical classes (covering all possible technical approaches for 3D displays) can be described using two fundamental physical criteria. Parameters of the displays are briefly analyzed with respect to the psychophysical properties of human 3D vision..

Key Words: 3D display, 3D Vision, stereoscopic displays, holography, volumetric displays

1 Introduction

3D displays cardinally differ from 2D displays by providing images carrying true information about the third dimension (about the depth) of 3D scenes. The "3D display" term here means "3D *information* display" which creates the spatial light distribution recognized by human vision as a 3D scene with *proper information quality*, i.e. with proper resolution, color and other important parameters of the image corresponding ideally to those of state-of-art 2D displays.

The keystone of any 3D display is an electronically addressed *work medium* – the *dynamic* (reversible) medium which modulates parameters of light in real time according to the amplitudes of input electronic signals to form spatial light distributions representing (simulating) the 3D scene. *Work space* is the space occupied by the work medium (by spatial light distribution appearing within the work medium). Information light distributions can be created inside the medium with the help of any suitable physical means such as light refraction, absorption, diffraction, scattering or generation.

2 Physical Criteria

1.1 CHARACTER Of Light Distribution Capable To Represent 3D Scenes

Any object of a 3D scene can be considered as a set **S** (Fig.**1**) of point light sources generating the light distribution **V** which propagates (in the form of optical waves) through a volume of free space (with finite physical width Z) and reaches the eyes in the form of light distribution **V'** (differing from **V** due to the properties of free space volume considered as an optical element with a certain transfer function). Human 3D vision perceives the light distribution **V'**, and the final 3D scene is created in human consciousness (brain).

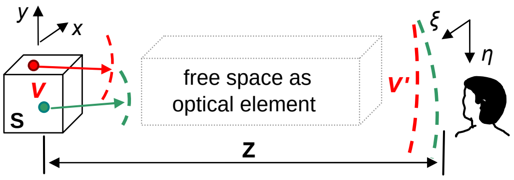


Fig.1 Optical sheme of natural 3D scene viewing

Light distribution $V'(\xi,\eta)$ perceived by human 3D vision is mathematically expressed by the Fresnel-Kirchhoff diffraction integral

$$V'(\xi,\eta)\cong\frac{\exp(ikZ)}{Z}\iint_{\infty}V(x,y)\exp\{\frac{ik}{2Z}[(x-\xi)^2+(y-\eta\}^2]\}dxdx \qquad (1)$$

Consider an elementary 3D object consisting of a single point light source (with unit energy) located at (x_0,y_0). The corresponding light distribution $V(x,y)$ is expressed by a delta function $\delta(x_0,y_0)$

$$\delta(x,y)=\infty \text{ at } x=x_0\,;\quad \delta(x,y)=0 \text{ at } x\neq x_0\,;\quad \int_{-\infty}^{+\infty}\delta(x,y)=1 \qquad (2)$$

Substituting (2) in (1) gives the light distribution $V_0'(\xi,\eta)$ input to the eyes in case of the point object light source

$$V_o'(\xi,\eta)\cong\frac{\exp(ikZ)}{Z}\exp[ik\left(\frac{x_o^2\xi^2+y_o^2\eta^2}{2Z}\right)] \qquad (3)$$

Expression (3) describes a spherical light wave whose radius of curvature carries information about the distance Z between observer and point light source. The intensity of the spherical wave is proportional to the intensity of the light source and inversely proportional to the distance Z. The angle of the spherical wave inclination (relative to the position of the eyes optical axes) is defined by coordinates (x_0,y_0) of the point source.

Let us consider how human vision distinguishes the depth of a 3D scene.

1.2 Basic Properties Of Human 3D Vision

Inside the human optical "apparatus" **there is no *direct* three-coordinate analyzer** of light distribution. In fact any incoming light distribution $V_0'(\xi,\eta)$ is always analyzed in two eyes via two 2D ("flat") light distributions formed on the retinas (Fig.**2**). Regardless of the initial 3D scene, the eyes

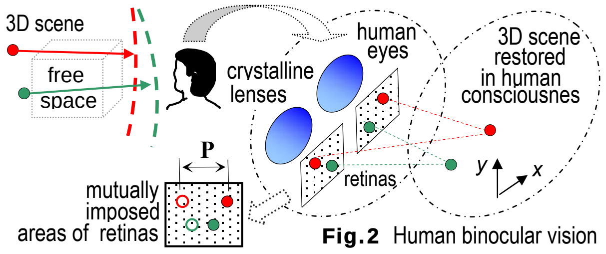


Fig.2 Human binocular vision

always receive spherical optical waves created by diffraction in a volume Z of free space. The crystalline lenses of the eyes concentrate the energy of each elementary spherical wave to corresponding points on both retinas representing two differing 2D projections of the 3D scene. The virtual 3D scene is formed in human consciousness mainly via binocular vision. **Binocular vision** synthesizes depth from the parallax P between the positions of a pair of points (on two retinas) corresponding to the pair of projections of the same elementary point light source. In the left bottom corner of Fig.**2** the parallax P is shown on corresponding points of two retinas.

THE FIRST CLASSIFICATION CRITERION

Human 3D vision can correctly reconstruct the information from a 3D scene if the work medium gives output light distribution in any of two possible physical forms which we define as ■ **object presentation** and

■ **Fresnel-diffraction** (or **spherical-wave**) **presentation**.

The **object** presentation (Fig.**3**) is a *point-by-point presentation* of scene objects in medium **M** described by $V(x,y)$. The **Fresnel-diffraction** described mathematically by $V(\xi,\eta)$ in formula (1) corresponds physically to a **spatial angular spectrum of spherical-waves** originating from the 3D scene object points (Fig.**4**).

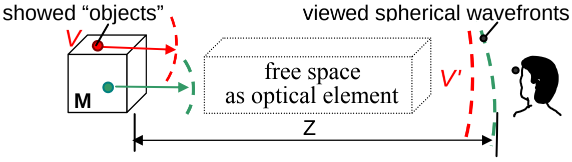


Fig.3 *Object* light distribution inside work (presentation) medium

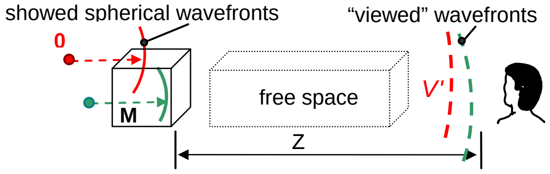


Fig.4 *Spherical-wave* distribution inside work (presentation) medium M

Now we consider the work medium **M** dimensionality (number of spatial coordinates) required to present a 3D scene. It is obvious that a three-coordinate work medium is able to present all three coordinates. Taking into account properties of binocular vision (Fig. **2**) it is clear that a pair of flat projections of a 3D scene provide the necessary information about all three coordinates of this 3D scene viewed at a certain angle. The work **medium can be three-coordinate or two-coordinate (3D or 2D medium).**

Consider whether we should include **1D** presentation medium as a separate case. **1D** light distribution itself can't be processed as a 3D scene. In fact, vision has a short **time-integrating memory** permitting storage of a whole 2D picture from sequentially incoming 1D light distributions. The minimal frequency F_{min} (permitting time sequential integration without flickering) depends on value of light exposure of each elementary area of the retina E according to the empirical formula:

$$F_{min} = 9{,}6\,(\lg E+2{,}3), \qquad (4)$$

where E – light exposure in lux.

So time-sequential **1D** light distributions always have equivalent **2D** distributions directly on the retinas (or indirectly - on intermeduate screen having a short-time integrating memory for storing 2D pictures from 1D light distributions).

THE SECOND CLASSIFICATION CRITERION

It is worth to consider only cases of **2D** and **3D** work medium (as 1D mediums are always used to form 2D light distributions via scanning in a 2D work space).

Comfortable viewing. It is highly desirable to provide not only the proper information content but also the conditions for comfortable viewing of the displayed 3D scenes (ideally as comfortable as viewing natural 3D scenes). Two critical basic psychophysical properties of human binocular vision are **accommodation** and **convergence**. Accommodation is focusing of the crystalline lenses on object-of-interest by changing the curvature with the eye muscles. Convergence is the intersection of the optical axes of both eyes on the object performed by mutual inward rotation. In perception of natural 3D scenes this is always a **match** between accommodation and convergence (Fig.**5**) because surface **1** of the focusing map of the crystalline lenses and the intersections of optical axes **2** and **3** of the eyes take place in one and the same point **0** of the viewed object. Mismatch between accommodation and convergence can cause visual discomfort.

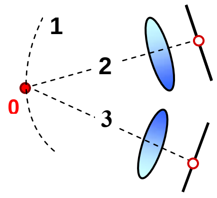


Fig.5 Accommodation and convergence

Conditions for high fidelity of 3D scene presentation..

Eye resolution and angle characteristics of 3D perception.

The eye resolution cue is shown in Fig.**6**. These parameters should be taken into account for estimation of required resolution of presented 3D images. The full field-of-view (FOV) of human vision is 180^0x135^0. The FOV of 3D binocular vision is 120^0x135^0 (120^0 is common horizontal 3D FOV for both eyes). The size of the pupil varies from 2 to 8 mm (depending on light exposure).

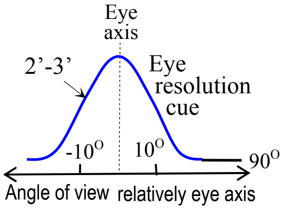


Fig.6 Eye resolution cue

An interesting psychophysical property is the effect of eye *saccades* (Fig.**7**) – fast involuntary micro movements of the eyes leading to continual fast micro changes of the viewing angle . The main purpose of saccades seems to be to permit biochemical recovery of the visual cells because it has been shown that steady light distribution causes the loss of their sensitivity to light. An important function of saccades in stereoscopic vision may be their possible contribution to naturalness or vividness of 3D

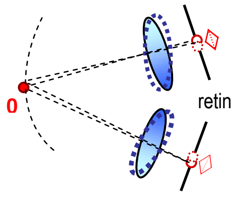


Fig.7 Eye saccades

scenes even for a static position of a viewer (due to continual slight changes in viewing angle by saccades).

There are some additional properties of human 3D vision. High fidelity 3D imaging must involve maximum quantity of mentioned properties of human 3D vision.

3 Technical Features Of Each Physical Class of 3D Displays

Corresponding to the two physical criteria there are **4 physical classes** of 3D displays:

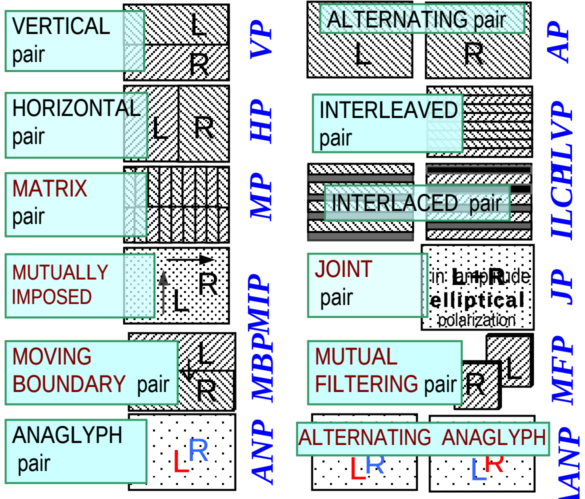
Object Presentation In 2D Medium
Object Presentation In 3D Medium
Fresnel-Diffraction (Spherical-Wave) Presentation In 2D Medium
Fresnel-Diffraction (Spherical-Wave) Presentation In 3D Medium.

Next we will consider the features of each class.

3.1 The Class "Object Presentation In 2D Medium (STEREOSCOPIC Displays)

If, with the help of a 2D screen, left **L** and the right **R** views of a 3D scene are directly presented separately to the left and right eyes (Fig.**8**) then human consciousness (the brain) restores this 3D scene at the corresponding angle of view. This class contains well-known stereoscopic displays. Stereoscopic displays are subclassified in accordance with their *stereoscopic formats* (Table **1**) - positions of **L** and **R** views on 2D screens permit presentation of these views separately to the left and right eyes. Brown designates the names of stereoscopic formats introduced here by the author and Blue - his abbreviations.

Table 1. Stereoscopic formats (names and abbreviations)



VERTICAL pair	VP	ALTERNATING pair	AP
HORIZONTAL pair	HP	INTERLEAVED pair	LVP
MATRIX pair	MP	INTERLACED pair	ILCH
MUTUALLY IMPOSED	MIP	JOINT pair	JP
MOVING BOUNDARY pair	MBP	MUTUAL FILTERING pair	MFP
ANAGLYPH pair	ANP	ALTERNATING ANAGLYPH	AANP

The presence of two adjacent pictures (in a cell of Table **1**) means a different topology (position) of L-R views in two adjacent (in time) frames (or fields) of an image on the screen of a 2D display.

Advantages of the class: ☐ highest compatibility with *standard* 2D electronic displays, information sources (cameras, computers) and with information transmission channels; ☐ highest functional flexibility, including the possibility of making panoramic 3D displays with immersive and adaptive capabilities (such as combining a high-resolution central image area with a low-resolution peripheral area to minimize bandwidth without loss of image quality owing to features of vision).

Disadvantages: ☐ a discrepancy (Fig.**9**) between accommodation **1** and convergence **2**, **3** causing visual discomfort. To minimize the discrepancy it is necessary to increase the distance between a viewer and the 2D screen, to minimize parallax value P, or to introduce correcting optical schemes.

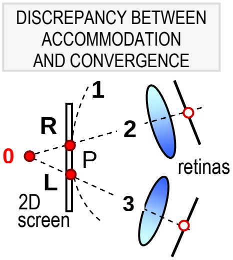


Fig.9 Main disadvantage of stereoscopic displays

3.2 The Class "Object Presentation in 3D Medium" (VOLUMETRIC Displays)

Only displays with *time sequential isotropic* optical modulation (Fig.**10**) are realized in practice now; in one version - using a mechanically rotating mirror, in another version - using a electrically switched multiple planes as a 3D display medium (Fig.**11**). In both versions **S** is the source of time sequential 2D images.

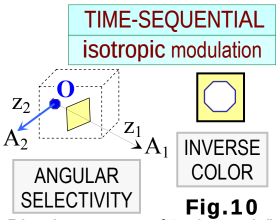


Fig.10 Disadvantages of "volumetric"

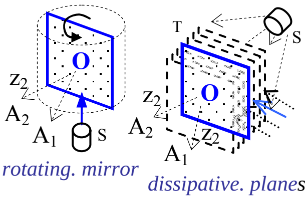


Fig.11 Practical "volumetric"

There is a time sequential light modulation here giving mutually independent partial light fluxes on all points of the display medium disposed along z-distance (z_1, z_2) relative to positions A_1, A_2 of the viewer. It leads to independent presentation of each point source on the retinas, this, in turn, causes a strong angular selectivity of the image relative to the observer's position (i.e. correct imaging of an arbitrary 3D scene can achieved only for a single angular position of the observer (for which the computer has calculated the occlusion for all points that should be hidden if the real objects were seen). All other positions of an observer lead to absence of proper occlusion and to the perception of points that should be hidden Also the time sequential method leads here to wrong additive mixing of colors on retinas of the viewer (instead of the correct subtractive mixing for sequentially disposed objects along z-axis as it takes place in real 3D scenes). As a result, inverse colors are produced. For example, the *blue* of dot **O** *is added* to the *yellow* of the rectangle, giving inversed *(white)* color in the area of intersection of dot and rectangle (instead of *subtracting* yellow from blue giving *black*).

Advantage of this class: ☐ accommodation is matched with convergence. **Disadvantages of time sequential versions of "volumetric" displays:** ☐ due to angular selectivity it is possible to a make walk-around image only for "wire frame" or high-transparency objects (where occlusion is not required); ☐ extreme high bandwidth (hundreds of Gbits/s) in case of even standard resolution 3D images, ☐ inverse color characteristics, ☐ impossible to represent natural scenes directly because of the absence of 3D optical volumetric pickups with electronic output (only computer-generated 3D scenes are feasible), ☐ impossibility of presenting panoramic images.

Using *nonisotropic* optical modulation in "volumetric" methods to overcome the problem of angular selectivity is problematic because of the extreme complexity of the resulting hypothetical 3D display device. Perfect object 3D modeling of a 3D scene can theoretically be created by using *time-*

overlapping (or simultaneous) optical modulation (Fig.**12**) in each z-layer. For example, one could use a multiplicity of parallel 2D displays with a synchronized addressing scheme. However, such a design is probably unrealizable in the near future because a stack of several hundred flat displays have great light losses, retro reflections, low reliability, and it is also problematic to provide isotropic light modulation.

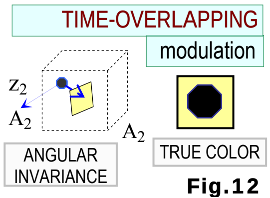


Fig.12 Perfect unrealizable “volumetric”

3.3 The Classes “Fresnel-Diffraction Presentation In 2D And 3D Display Medium”

Equation (3) describes the complex envelope $V_0'(\xi,\eta)$ of an optical spherical wave in the form of amplitude and phase modulation of an optical carrier (even if the object presentation $V(x,y)$ is a real amplitude function). The optical carrier is omitted from the notation (3) because it is the same multiplier for both parts of (3) since it has one and the same value in all points of space where the optical wave propagates. The physical presence of an optical carrier with extremely high frequency prevents direct recording of traveling optical waves in any dynamic medium. Now the only practical way to record complex content of an optical wave is excluding or greatly decreasing the frequency of the optical carrier by employing a coherent reference wave having the same or nearly the same frequency (it is the essentially the well known dynamic holography method).

Fresnel zone pattern (Fig.**13**) shows how amplitude and phase variations can be recorded in the form of a real amplitude function (in the form of interference fringes that are the “ink” of the true holographic record). This is an elementary Gabor hologram.

Methods of *linear optics dynamic holography* and *nonlinear optics dynamic holography* should be differentiated. Linear and nonlinear optics correspond to linear or nonlinear dependence of electrical polarization of display medium on the electrical field value of the optical wave.

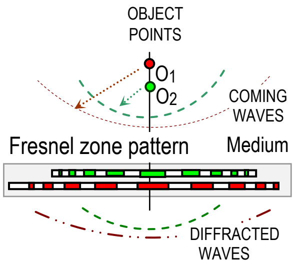


Fig.13 Interference frnges

3.3.1 Dynamic Interferometric Recording of Spherical-Waves (HOLOGRAPHIC DISPLAYS with Thin and Thick Dynamic Fresnel Holograms)

Each point of a 3D object is represented here in the form of a spherical wave that is recorded in the structure of a **dynamic** interferometric fringe pattern in this way: the visibility of interferometric fringes corresponds to the brightness of the perceived object points; the spatial frequency of fringes - to angular position of the object points; and the curvature of the fringes - to the distance from the points.

Only acoustooptic modulators (AOM) can be used now for synthesizing such dynamic holograms from electronic information signals. Raman-Nath diffraction AOM is the traveling-wave analog of Leith-Upatnieks thin off-axis hologram (Fig.**14**). Bragg diffraction AOM is a traveling-wave analog of Denisjuk thick holograms (Fig.**15**). Each fringe of the Fresnel zone pattern is modulated by a spatial grating with period **p** to attain the angular separation of the desired 3D image. In practice only Bragg-diffraction AOM are feasible because of their ability to work with broadband (nonlaser) light sources.

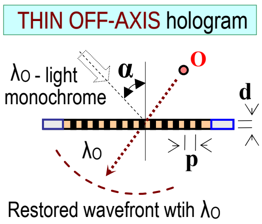


Fig.14 Leith-Upatnieks hologram

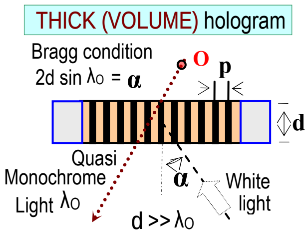


Fig.15 Real holographic medium

The size of the eye’s pupil defines the physical limit of required angular resolution of holographic views according to the Rayleigh criteria.

Advantages of holographic displays with thick dynamic holograms: ☐ multiview capability with such angular resolution that it is possible to use the effect of eye saccades for increasing realism ☐ accommodation is matched with convergence because the eyes can’t focus on the surface of a hologram ☐ it is possible to build panoramic displays.

Disadvantages: ☐ one-dimensionality and a small aperture of high-frequency (10-20 Hz) AOM leads to cumbersome optical layouts, to the limits of spatial resolution of the resulting equivalent 2D or 3D dynamic holographic medium, ☐ presently, there is no another real-time holographic medium with the required resolution of several thousand lines per millimeter (necessary for high-grade holographic record).

3.3.2 3D Scene Wavefront Creation By Nonlinear Optics Dynamic Holography

Nonlinear optics can be used for very high amplitudes of electrical vector of the optical wave (comparable to the electrical fields of the display mediums atomic structure). Wavefronts can be created, for example, by four-wave mixing . **Disadvantages:** ☐ low efficiency (about several percents) of real-time nonlinear optical transformations lead to poor 3D images, ☐ very high electrical fields of the optical wave conducts to low reliability, to raised danger of display breaking and to difficulty of making a compact arrangement of the information carrying pixels. There are no serious prototypes of such 3D displays (only 3D visual effects have been shown).

4 Conclusion

The table **2** summarizes short-term physical names (modalities) of 3D display technical approaches:

Table 2. Names of 3D displays

Name as **physical definition**	**Habitual** name
Object 2D-medium (O2D) display	Stereoscopic display
Object 3D- medium (O3D) display	Volumetric display
Fresnel-diffraction 2D-medium (F2D) display	Holographic 2D-medium display
Fresnel-diffraction 3D-medium (F3D) display	Holographic 3D-medium display

3D display designs of any type can be analyzed into four technical modalities - **O2D**, **O3D**, **F2D**, **F3D** whose inherent technical features (briefly considered in paragraph 3) define all the features of the whole 3D display design.