Janusz Konrad and Michael Halle

3-D Displays and Signal Processing

An answer to 3-D ills?

Three-dimensional (3-D) perception is an intrinsic part of the human experience. While most people gain the majority of their spatial information through vision, and approximately 90% of the population benefit from stereopsis, display systems have historically reproduced only two-dimensional depth cues. Over the last 150 years, many attempts have been made to exploit stereopsis in various 3-D displays; while several achieved limited commercial success, none have attained equal status to their 2-D counterparts. Today, novel electronic display technologies, powerful microprocessors, and advanced signal processing algorithms are about to open a new era for 3-D displays. Signal processing specifically focused on 3-D imaging will, in large part, determine the viability of these emerging 3-D display systems.

In this article, we overview today's main electronic 3-D display technologies from a signal processing perspective. We describe the underlying physics, and point out benefits and deficiencies of various displays. We discuss the general role of signal processing and provide specific

Digital Object Identifier 10.1109/MSP.2007.905706

examples of signal processing helping address certain display deficiencies. We highlight challenges awaiting signal processing in quest of the ultimate 3-D experience.

INTRODUCTION

Virtual presentation of the true 3-D world has fascinated humanity for centuries. Seeking a "being there" experience, we have developed various visual devices capable of mimicking the appearance of a 3-D object or scene. Such devices use mechanical, optical or electro-optical means to simulate a field of light that produces a different projection on the retinas of the viewer's two eyes. The brain interprets the disparity of these light patterns as depth. This process is known as *binocular stereopsis*.

It is important to realize that retinal disparity is not the sole cue human visual system relies on to perceive depth; indeed, it is often not even the strongest cue. Two-dimensional cues such as perspective, occlusion, relative object size are critically important for depth perception; motion parallax helps us understand the structure of objects that we can move around (or that move in front of us). In this paper, however, we limit our discussion to displays that deliver binocular stereopsis and, when applicable, motion parallax, and include other cues only in relationship to them.

Different 3-D display technologies offer distinct benefits but are also plagued by unique deficiencies. From the early days of stereoscope (mid 19th century), through parallax stereograms (early 20th century), polarized 3-D movies, personal stereo cameras (1950s), and holography (1960s and 1970s), to quality 3-D films by IMAX^{TM} (1980s and 1990s) and electronic 3-D (today), advances in materials, electronics and optics have lead to significant improvements in 3-D image quality, visual comfort and benefit/cost ratio. Due to these improvements, 3-D displays have become an important tool in a variety of specialized applications, including image-guided surgical procedures, remote guidance of robots at dangerous work sites, battlefield reconnaissance, and scientific visualization.

The adoption of 3-D displays is strongly driven by the ongoing digital multimedia revolution. While 3-D imaging of previous decades relied on custom components and technologies far outside the mainstream, 3-D display devices of today can take advantage of an all-digital content handling chain that includes capture, processing, editing, and display. Of particular interest to the signal processing community, 3-D aware algorithms can be added to this existing pipeline in a natural way without reinventing all of the other steps. This new compatibility allows for emerging opportunities in 3-D specific processing, including compression, anti-aliasing, and image enhancement. We believe that this new, exciting, and largely unchartered area can become a highly productive "playground" for digital signal processing as 3-D becomes a larger part of the imaging pipeline.

The article is divided into four parts. First, we describe *stereoscopic displays*, the simplest and most common systems that provide binocular stereopsis to the viewer. Next, we discuss *planar multiview displays*, which provide "look around" by displaying multiple viewpoints of the scene. *Holographic displays* and electroholographic systems are described as a distinct approach to planar 3-D display using diffractive optical elements. Finally, we discuss *volumetric 3-D displays*, which form a space-filling image volume using a variety of scanning techniques.

While this article provides some historical context for different 3-D display technologies, more general references such as Okoshi [1] and Benton [2] offer the reader more extensive and complementary background material.

[VIRTUAL PRESENTATION OF THE TRUE 3-D WORLD HAS FASCINATED HUMANITY FOR CENTURIES.]

PLANAR STEREOSCOPIC DISPLAYS

The earliest attempts at stereoscopic 3-D involved placing two drawings in front of viewer eyes and separating them by a mechanical barrier. Such a device, known as the *stereoscope* was invented by Charles Wheatstone and improved by David Brewster and others in mid 19th century [3]. While Wheatstone used line drawings, photographs from two closely-spaced cameras soon became the dominant 3-D content. Stereoscopes became extremely popular during the Victorian era, and many different variants of the technology emerged before George Eastman ushered in the era of personal, monoscopic photography. Since that time, 3-D photography has never rivaled the popularity of its "flat" counterpart.

A stereoscope delivers two 2-D images through partitioned optical channels to a viewer's eyes, a principle of operation shared by all stereoscopic displays. Where display types differ is in their method of partitioning. The stereoscope, for example, uses two separate lenses and two printed or electronic images. Head-mounted or goggle-based electronic display systems use a similar channel-partitioning technique. We will not consider these technologies further. Rather, we will concentrate on two other ways to multiplex the left/right components of a stereoscopic image: glasses-based systems and autostereoscopic, or glasses-free, technologies.

STEREOSCOPIC DISPLAYS WITH GLASSES

Glasses-based stereoscopic imaging uses filters or shutters, one for each eye of each viewer, to present left and right channels of view information from a single display or screen. Glasses-based systems have the great advantage of being able to provide an entire audience with a 3-D imaging experience with reasonable economy. While film-based cinema was the original motivation for the development of glasses-based stereo, these technologies are applicable to both desktop and large-scale electronic 3-D displays today.

WAVELENGTH-DIVISION MULTIPLEXING

The introduction of color in printing and photography, and now electronic displays, provides a natural mechanism for left/right view channel separation based on spectrum division. The simplest division of the visible light spectrum (370–730 nm) is into two bands, for example 370–550 nm and 550–730 nm. Since ideal "brick-wall" optical filters cannot be realized in practice, typical transmission characteristics are similar to those shown in Figure 1(a), that correspond to red (left) and blue (right) lenses. The basic *anaglyph* method places the left-image data in the red channel of a color photograph or electronic display, and the right-image data in the blue channel. Although there exists some flexibility in composing the green channel, this strongly depends on spectral characteristics of the glasses. For example, with red/blue or red/cyan glasses placing the right-image data in the green channel leads to much improved depth perception over placing of the left-image data there since blue-green or cyan-green mixtures fall in the passband of the right spectacle unlike the red-green mixtures that do not match well the left lens.

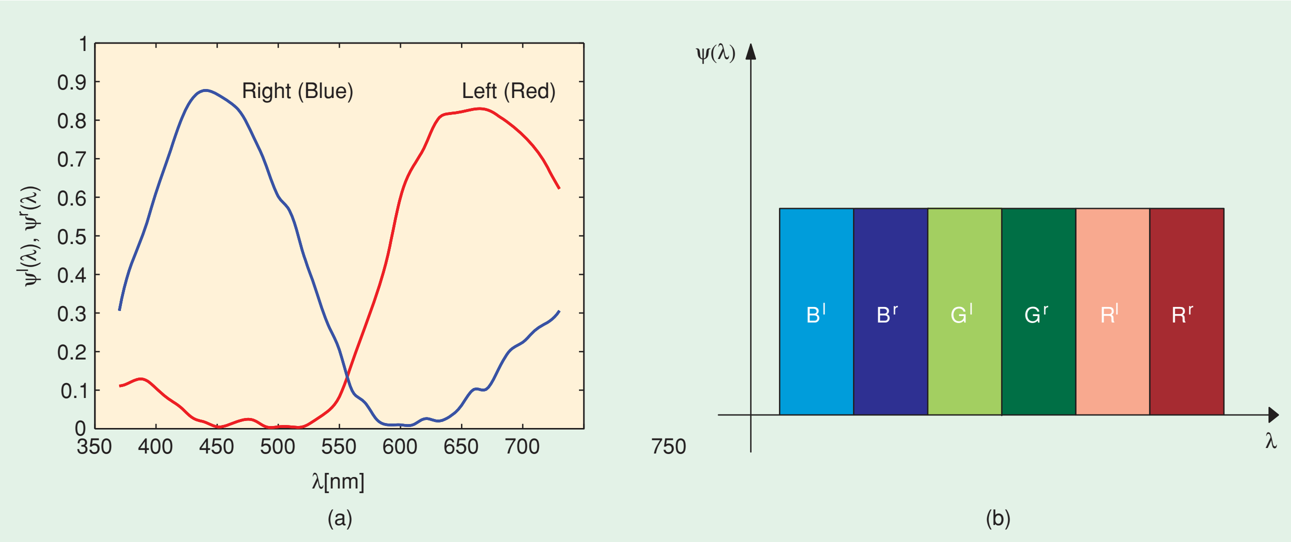
Historically, anaglyph 3-D reproduction was developed in mid 19th century for line drawings and then extended to photographs; it retains these uses today as the most printer-compatible 3-D system. It gained popularity with the advent of motion pictures in the early 20th century, but was mostly subsumed by polarized systems by the 1950's. Anaglyphic 3-D resurged after the color CRT (cathode ray tube) became commercially available; in 1960s some TV programs and reprinted 3-D movies were broadcast in anaglyph mode.

The most important benefit of anaglyph stereo is that it can be used on essentially any device capable of reproducing color, whether electronic, film or paper. Additionally, with simple, easy-to-manufacture glasses it is the least expensive technique for mass 3-D visualization. Its main deficiencies, however, include visual discomfort due to different spectral content presented to each eye (color rivalry) and optical crosstalk between channels (poor spectral separation of colored lenses). Anaglyphs also have limited color reproduction capability.

Although the channel crosstalk and color gamut can be, to a degree, controlled by a careful selection of color lenses vis-à-vis screen phosphors or print dyes, the range of possible improvements is rather small. This is primarily due to a limited selection of spectral lens characteristics and rather rudimentary algorithms for anaglyph preparation. Significant improvements, however, are possible by employing signal processing methods to optimize anaglyph images with respect to reproducible color gamut given spectral absorption curves of the lenses, spectral density functions of the display primaries, and colorimetric characteristics of the human visual system. A formulation of this problem has been recently proposed together with a projection-based solution [4] (see "Breathing Life into Anaglyph 3-D"). The method produces distinctly improved color gamut of the perceived 3-D images in comparison with rudimentary algorithms. This technique is a good example of signal processing methods helping enhance 3-D display technologies.

MULTI-BAND FILTERS

The main deficiency of the anaglyph approach stems from the fact that light captured by each eye is concentrated in a single wavelength band (e.g., 370–550 nm or 550–730 nm), thus preventing full color perception in each eye and inducing color rivalry between eyes. Additionally, poor spectral separation characteristics of the lenses induce optical crosstalk. An improvement over the anaglyph method is multi-band wavelength-division multiplexing, where the visible light spectrum is divided into two complementary *sets of wavelength intervals*. When those sets are uniformly spread out across the



[FIG1] Spectral transmission curves of: (a) practical anaglyph (dual-band) Roscolux filters: orange-red (#25) and brilliant-blue (#69), and (b) ideal multi-band filters illustrating the principle of operation of the Infitec GmbH 3-D display system.

light spectrum (Figure 1(b)), a vastly improved color reproduction in each eye is possible. Recently, Infitec GmbH, Germany developed a pair of tri-band filters (coarse comb filters) [6] whose spectral responses are complementary but include a sufficient composition of wavelengths to warrant good color reproduction in each eye.

For 3-D visualization, two projectors are equipped with complementary filters while viewers wear corresponding glasses. With suitable processing, colors are represented in terms of spectral characteristics of left and right filters independently, i.e., in terms of left-lens primaries $R^lB^lG^l$ and right-lens primaries $R^rG^rB^r$ (Figure 1(b)). If this color representation is accurate, no visual discomfort results, despite different color primaries used for the left and right eyes, because the human eye cannot distinguish different compositions of the same color. Two additional benefits of this technology are the facts that filters with sufficiently accurate spectral responses to minimize optical crosstalk can be manufactured today, and that nondepolarizing screens are not needed.

The perceived 3-D image quality in such a system depends on the accuracy of color representation vis-à-vis filter spectral functions, and the complementarity of these functions between the left and right channels (amount of crosstalk). While the first issue is a standard problem in colorimetry, the latter one can be addressed by techniques similar to that discussed in "Dealing with Ghosts."

Note, that multi-band filters are relatively new to the commercial market and are significantly more expensive than either traditional anaglyphic or polarized glasses. At this time, the majority of multi-band filter systems require two projectors, although as of late 2006 Barco offers a single-projector Infitec-based model.

LIGHT POLARIZATION

Polarization multiplexing offers another method for designing glasses-based 3-D systems, permitting the kind of full-color reproduction not possible with anaglyph systems. In a typical arrangement, two projectors are equipped with differently-polarized filters. For example, if linear polarization is used, one projector is equipped with a horizontally-oriented polarizing filter, the other with a vertical polarizing one. A special nondepolarizing screen is necessary to ensure that polarization is maintained during projection. When a

BREATHING LIFE INTO ANAGLYPH 3-D

The perceived 3-D image quality in anaglyphic viewing can be improved by means of signal processing as proposed by Dubois [4]. The main idea is to optimize the anaglyph image so that displayed on a screen with known spectral characteristics and viewed through lenses with known absorption functions, the image is perceived as close as possible (in sense of a proposed metric) to the original stereo pair presented without glasses.

Let $\{I_i^l, I_i^r\}$, $i = 1, 2, 3$, be an RGB stereoscopic image pair. A pixel at location x is described by two RGB triplets, expressed as vectors: $\boldsymbol{I}^l[x] = [I_1^l[x], I_2^l[x], I_3^l[x]]^T$ and $\boldsymbol{I}^r[x] = [I_1^r[x], I_2^r[x], I_3^r[x]]^T$. Note that $I^l[x]$ and $I^r[x]$ are triplets of tristimulus values expressed with respect to some primary system that may be gamma-corrected, so that they can be directly displayed on a standard CRT monitor.

If $s_i(\lambda)$, $i = 1, 2, 3$, are spectral density functions of the RGB display phosphors, and $\bar{p}_k(\lambda)$, $k = 1, 2, 3$, are color-matching functions for selected primary colors, then the left- and right-image colors perceived at x can be expressed as follows:

$$\widetilde{I}_k^l x = \sum_{j=1}^{3} c_{kj} I_j^l[x], \quad \widetilde{I}_k^r x = \sum_{j=1}^{3} c_{kj} I_j^r[x],$$
$$c_{kj} = \int \bar{p}_k(\lambda) s_j(\lambda) d\lambda, \quad k, j = 1, 2, 3. \tag{1}$$

In vector notation, we can write: $\widetilde{\boldsymbol{I}}^l[x] = \boldsymbol{C}\boldsymbol{I}^l[x]$ and $\widetilde{\boldsymbol{I}}^r[x] = \boldsymbol{C}\boldsymbol{I}^r[x]$, where $[C]_{kj} = c_{kj}$. This is a simple transformation between two sets of primary colors. For a specific monitor, this results in a fixed 3×3 matrix C (see [4] for an example). At each x, the sextuplet of color coordinates $\widetilde{\boldsymbol{I}}[x]$ can be expressed in terms of the original sextuplet $I[x]$ as follows $\widetilde{\boldsymbol{I}}[x] = [(\widetilde{\boldsymbol{I}}^l[x])^T (\widetilde{\boldsymbol{I}}^r[x])^T]^T = C_2\boldsymbol{I}[x]$, where the 6×6 matrix C_2 consists of two matrices C on the diagonal and zeros elsewhere.

Let I_j^a, $j = 1, 2, 3$, be an anaglyph image to be displayed on the same monitor but viewed through glasses with colored lenses with spectral transmission functions $\psi^l(\lambda)$ and $\psi^r(\lambda)$ (Figure 1(a)). A similar transformation of color coordinates takes place again except for the presence of color filters. The two color transformation matrices are:

$$[A^l]_{kj} = a_{kj}^l = \int \bar{p}_k(\lambda) s_j(\lambda) \psi^l(\lambda) d\lambda,$$
$$[A^r]_{kj} = a_{kj}^r = \int \bar{p}_k(\lambda) s_j(\lambda) \psi^r(\lambda) d\lambda \tag{2}$$

and new color coordinates are $\widetilde{\boldsymbol{J}}^l[x] = A^l\boldsymbol{I}^a[x]$, and $\widetilde{\boldsymbol{J}}^r[x] = A^r\boldsymbol{I}^a[x]$. Combining the left and right color coordinates into a sextuplet, leads to a simple transformation $\widetilde{\boldsymbol{J}}[x] = R\boldsymbol{I}^a[x]$, where $\boldsymbol{I}^a[x] = [I_1^a[x], I_2^a[x], I_3^a[x]]^T$ and $R^T = [(A^l)^T (A^r)^T]$ is 6×3 matrix.

The goal is now as follows. Given a stereo pair $\{\boldsymbol{I}^l[x], \boldsymbol{I}^r[x]\}$ with $0 \le I_j^l[x], I_j^r[x] \le 1$ for $j = 1, 2, 3$, find an anaglyphic image $\boldsymbol{I}^a[x]$ with $0 \le I_j^a[x] \le 1$ for $j = 1, 2, 3$, such that the image $\widetilde{\boldsymbol{J}}$ perceived through the glasses is as similar as possible to the input stereo pair $\widetilde{\boldsymbol{I}}$. Dubois proposed a metric to numerically compute the subjective difference between a stereo pair and anaglyphic image based on certain heuristics, and developed an optimization method based on projection. The images produced by this method result in a wider color gamut perceived by viewers compared to rudimentary anaglyph preparation methods. This method is currently considered state-of-the-art technique for anaglyph preparation [5].

viewer uses correspondingly-polarized glasses, each eye captures light from (ideally) one view only, producing a sensation of depth.

The ability of a polarized system to render a full color gamut greatly reduces the visual discomfort typical of anaglyphic systems. In real-world systems, however, crosstalk

DEALING WITH GHOSTS

The optical crosstalk perceived by a viewer in front of 3-D screen manifests itself as double edges or "ghosts" at high-contrast features misaligned between left and right images due to disparity. In dual-projector systems with circular polarization, the crosstalk is due to imperfect light extinction in the glasses and depolarizing properties of the screen. In systems using CRT monitor (or projector) and liquid-crystal shutters, the crosstalk is caused by phosphor persistence of the CRT (primarily green), imperfect light extinction of LCS in the opaque state (light leakage) and by timing errors of LCS (opening/closing too early or too late). In systems using a single DLP projector with LCS, phosphor persistence is not an issue and only extinction characteristics and timing errors of the shutters play a role. Although in each system improvements are possible by careful selection of components, manipulation of image contrast, and adjustment of disparity magnitude, the degree of improvement is quite limited.

A significant reduction of perceived crosstalk, even its complete elimination under certain conditions, is possible by employing signal processing. The basic idea is to create "anti-crosstalk," i.e., pre-distort images so that upon display ghosting is largely suppressed [8]. Recently, a computationally-efficient algorithm that accounts for specific screen and LCS characteristics has been developed [9]. The algorithm is based on a simple crosstalk model:

$$J_i^l = I_i^l + \phi_i\left(I_i^r, I_i^l\right), \quad J_i^r = I_i^r + \phi_i\left(I_i^l, I_i^r\right), \quad i = 1, 2, 3, \tag{3}$$

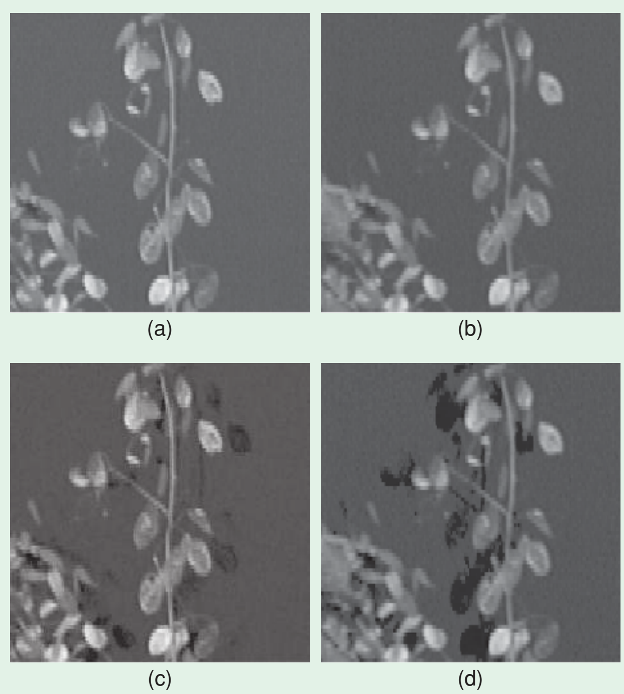
where J_i^l, J_i^r are RGB components of images perceived by the left and right eyes, respectively, I_i^l, I_i^r are RGB components driving the monitor, and ϕ_i's are *crosstalk functions* for the three color channels. The crosstalk functions ϕ_i quantify the amount of crosstalk seen by an eye in terms of unintended *and* intended stimuli. They are dependent on the particular CRT/LCS combination used and need to be quantified; example functions for a Sony Trinitron monitor and Stereographics Inc.'s CrystalEyes™ glasses can be found in [9]. Note that the above crosstalk model ignores the point spread functions of the CRT and LCS.

If the mapping (3), that transforms (I_i^l, I_i^r) into (J_i^l, J_i^r) for $i = 1, 2, 3$, is denoted by $\mathcal{T}$ with the domain $\mathcal{D}(\mathcal{T})$ and range $\mathcal{R}(\mathcal{T})$, then the task is to find the inverse mapping $\mathcal{T}^{-1}$ that transforms (I_i^l, I_i^r), images we would like to see, into crosstalk-biased images (G_i^l, G_i^r) that actually drive the monitor, i.e., find (G^l, G^r) satisfying:

$$I_i^l = G_i^l + \phi_i\left(G_i^r, G_i^l\right), \quad I_i^r = G_i^r + \phi_i\left(G_i^l, G_i^r\right), \quad i = 1, 2, 3. \tag{4}$$

For given crosstalk functions ϕ_i, this mapping can be computed off-line and stored in a look-up table; for 8-bit color components this is a $256 \times 256 \times 3 \times 2$ table (less than 400 kB), while for 10-bit components it requires no more than 6.3 MB and can be easily handled by modern graphics cards.

Since for $\mathcal{D}(\mathcal{T}) = [0, 255] \times [0, 255]$, $\mathcal{R}(\mathcal{T})$ is only a subset of $[0, 255] \times [0, 255]$, the inverse mapping $\mathcal{T}^{-1}$ operating on $[0, 255] \times [0, 255]$ may result in negative tristimulus values, that cannot be displayed. The algorithm is trying to "carve out" intensity notches (Figure 2), that will get filled with the unintended light, but in dark parts of an image this is not possible. Two solutions preventing negative tristimulus values are: linear mapping of RGB components, at the cost of reduced image contrast, and black-level saturation, that leads to loss of detail in dark image areas ("black crush"), both applied prior to $\mathcal{T}^{-1}$. Since neither solution is acceptable, a compromise between crosstalk reduction and image quality degradation is usually sought [9]. This method has been successfully used in LCS-based systems but it is equally applicable to polarization-based systems using CRT monitors, CRT projectors or DLP projectors. Example of application of the above algorithm is shown in Figure 2. Note the reduced image intensity in areas corresponding to bright features present in the other image; the unintended light due to optical crosstalk will fill those "holes" during display leading to crosstalk compensation.



[FIG2] Example of crosstalk compensation: (a) left, and (b) right original images (luminance only), and (c–d) the same images after crosstalk compensation using the algorithm described in [9] with luminance clipping below the amplitude of 30.

between left and right views in inevitable. View separation using linear polarizers is both imperfect and sensitive to rotation: as orientation of the glasses departs from that of projector polarizers, e.g., due to viewer head motion, the optical crosstalk between channels increases. These effects can be largely mitigated by employing circularly-polarized filters (clock-wise and counter clock-wise). Even when using circular polarizers, some optical crosstalk remains due to insufficient extinction properties of the filters, and light depolarization on the screen. In addition, circular polarizers are more expensive than linear ones.

The two-projector polarized systems have two further disadvantages. Polarization filters block a significant amount of light, so projectors with increased brightness are needed. Second, the alignment of the two projected images must be carefully maintained, since image misalignment is a major cause of eye strain.

THE 3-D DISPLAY DEVICES OF TODAY CAN TAKE ADVANTAGE OF AN ALL-DIGITAL CONTENT HANDLING CHAIN THAT INCLUDES CAPTURE, PROCESSING, EDITING, AND DISPLAY.

The projector alignment issue can be eliminated if static polarizers in front of left and right projectors are replaced by a dynamic polarization modulator in front of single projector. By interleaving left and right images in a video sequence, and switching the modulator polarity in synchronism with vertical synchronization impulses, the projected left and right images are differently polarized, and can be easily separated by viewer-worn glasses. However, a projector capable of doubled refresh rate is needed. Two examples of this technology are: *ZScreen*® developed by Stereographics Corp. (currently RealD Inc.) that circularly polarizes consecutive images and μ**Pol**™ (micropol) developed by Vrex Inc. which linearly polarizes consecutive lines of an image. The former has been deployed commercially in hundreds of electronic cinemas to enable 3-D movie projection. (This technology has lead to a recent revival of 3-D movies. The 2004 animated film *Polar Express*, the first movie released simultaneously in conventional and IMAX 3-D format, grossed 35% of its revenue from the 2% of the cinemas that chose to show the 3-D version. Polarized 3-D IMAX has been so successful, that conventional monoscopic IMAX movies are now often labeled “2-D” to avoid audience disappointment.)

Although the dynamic modulator solution does away with left/right image alignment issues, light output and crosstalk remain of concern. While the light absorption in a filter or modulator can be compensated for by increased brightness of a projector, the crosstalk problem necessitates signal processing solutions, such as those to be discussed in the next section; by pre-processing the left and right images one can largely suppress the perceived crosstalk. Such a solution is applicable to systems using circular polarization since crosstalk is largely shift-invariant. In systems with linear light polarization, the crosstalk is rotation-variant and the problem is rather intractable.

LIGHT SHUTTERING

The image alignment and light output issues can be resolved by employing shuttered glasses (e.g., liquid-crystal shutters) instead of polarized glasses. Rather than light polarization, light blocking is applied by means of fast switching lenses that, working in synchronism with the screen, become opaque when the unintended view is rendered on the screen, and transparent—when the intended view is displayed.

CRT displays in combination with liquid-crystal shutters (LCS) have long offered high quality stereoscopic visualization on the desktop. Such systems are characterized by full CRT spatial resolution, full CRT color gamut, and no flicker for CRTs with refresh rates of at least 120 Hz. However, this approach is relatively expensive since each viewer needs a pair of costly LCS glasses. Moreover, similarly to 3-D systems using polarization modulator and polarized glasses, the CRT/LCS combination suffers from optical crosstalk. The source of the crosstalk is different however, with the main factors being screen phosphor persistence, LCS extinction characteristics and LCS timing errors [7].

Similarly to 3-D displays using circular polarization, the CRT/LCS crosstalk is approximately shift-invariant; CRT phosphor persistence and LCS light extinction characteristics are quite uniform spatially (although the latter slightly changes towards shutter periphery). This approximate shift-invariance permits application of signal processing algorithms to pre-distort the original left and right images in order to compensate for the crosstalk added during display [8], [9]. Details of such an approach are discussed in “Dealing with ghosts”.

The widespread replacement of CRTs by slower-response LCD panels has made “stereo on the desktop” less convenient. Planar’s StereoMirror display [10] offers a desktop solution that is compatible with passive polarized glasses by combining two LCDs and a beamsplitter that flips the polarization of one of the panels. Image registration between the two views can be very good because of the flatness of the optical panels, although since one image is seen in reflection, it must be flipped left to right before display.

AUTOSTEREOSCOPIC DISPLAYS

The need to use glasses for viewing stereoscopic content has long been identified as a major inconvenience due to physical viewer discomfort, potential for projector misalignment, and the cost and care of glasses. Moreover, in scenarios where both 3-D and 2-D (nonsynchronized) screens are in the field of view of the glasses, liquid-crystal shutters optically interfere with nonsynchronized displays resulting in screen flicker and, therefore, additional viewer discomfort. Similarly, polarizing glasses interfere with viewing conventional LCD monitors.

As an alternative to systems using glasses, a number of techniques have been proposed since the early 1900's [1] that apply spatial multiplexing of left and right images in combination with a light-directing mechanism that presents those views to the viewer's eyes. These *autostereoscopic* techniques are directly applicable to electronic displays [11], [12]. The two major multiplexing techniques are *parallax barrier displays* (originally called parallax stereograms), and *microlens displays* (often called just lenticulars).

In parallax-barrier displays, an opaque layer (e.g., a very thin sheet of aluminum or inked screen) with narrow regularly-spaced slits is placed very close to a pixel-addressable screen, such as a plasma or LCD panel (Figure 3(a)). Note that for a given slit each eye's viewing angle is different and also that each slit acts horizontally as an approximate pin-hole projector. If the slits are precisely aligned with pixel columns, and the overall geometry is carefully adjusted (pixel pitch, slit width and pitch, barrier distance from the screen surface), a viewer sitting at a prescribed distance from the screen will see different sets of pixels with each eye, corresponding to the left and right views.

In lenticular displays, a sheet of narrow, thin cylindrical microlenses is typically attached to a pixel-addressable screen at approximately one focal length, so that light passing through lenses focuses the underlying panel (Figure 3(b)) to either infinity or the intended viewing distance. Again, the viewing angles for a given microlens are different and each microlens focuses light horizontally. If the microlenses are precisely aligned with pixel columns, and pixel pitch, microlens pitch and focal length are carefully selected, a viewer sitting at a prescribed distance will also see different sets of pixels with each eye.

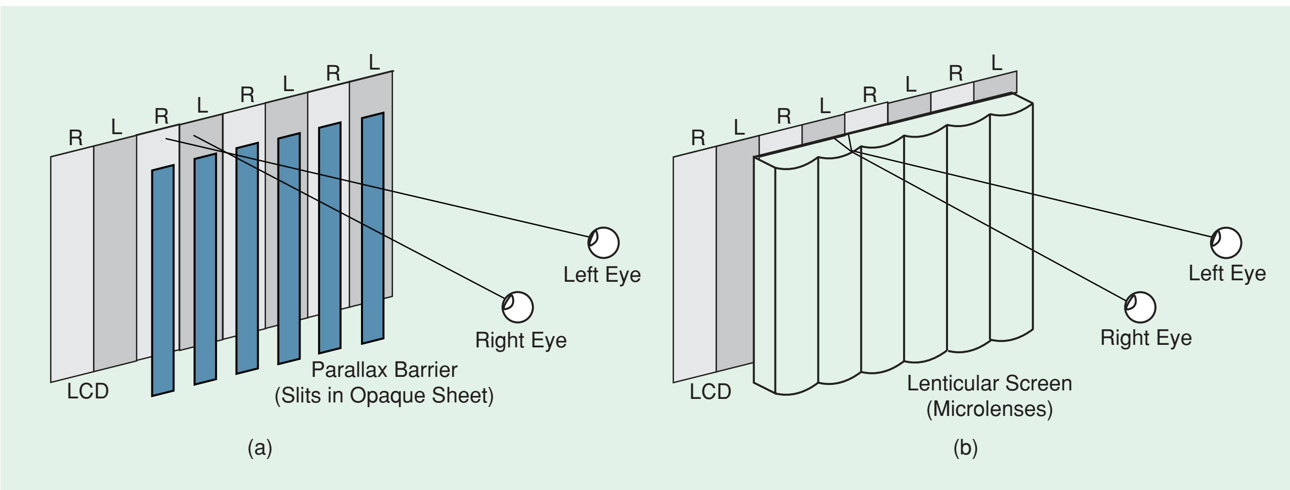
The locations in front of a screen from which a consistent 3-D perception is experienced, i.e., the left eye sees pixels with left-image data and the right eye sees pixels with right-image data, is called a *viewing zone*. The size of the viewing zones is related to the geometry of the display mentioned above.

When viewer's eyes straddle two different viewing zones in a two-view display, the presented left and right views are reversed and the viewer sees a *pseudoscopic*, or depth-inverted, image. The physical adjacency of correct (orthoscopic) and inverted (pseudoscopic) view zones can lead to visual confusion and discomfort for the viewer. Both parallax-barrier and lenticular technologies suffer from this shortcoming, which can be considered a type of optical crosstalk since one slit or lens can produce an image using its neighbor's data. Also, lens aberrations may contribute to crosstalk by reducing contrast of images from adjacent pixels.

One method of reducing crosstalk in parallax barrier displays is to reduce the width of each slit aperture. This approach has the disadvantage of reducing the display's light efficiency. Crosstalk can also be minimized by introducing guard bands with no image information in between the view data for each slit/microlens. With guard bands, the image would black out for one or both eyes rather than display neighbor's data, but only at the expense of reducing viewing zone size. In electronic displays guard bands are often impractical because of structure of the the underlying pixel array. In practice, display systems are designed to balance crosstalk, size of viewing zones and light output requirements, although the "sweet spot" of an autostereoscopic display, i.e., location where the image appears depth-correct, tends to be fairly narrow.

A few additional remarks are in order here. CRT displays cannot be used with parallax-barrier or lenticular technology because a precise geometric relationship between pixels and slits/microlenses is required; light patterns on a CRT screen suffer from unacceptable jitter. While parallax-barrier systems are easier to manufacture than lenticulars, their main deficiency is a reduced light output due to the nontransparent barrier.

Finally, both spatially-multiplexing techniques suffer from the loss of horizontal resolution delivered to each eye; half of the pixels in each horizontal row are delivered to left and half—to right eye. In order to spatially-multiplex left and right images on the screen, both images need to be horizontally



[FIG3] **Schematic illustration of how (a) parallax-barrier and (b) lenticular displays work.**

sub-sampled by the factor of two. Clearly, such sub-sampling necessitates prior half-band horizontal lowpass filtering in order to avoid aliasing. Such filtering is a special case of more general pre-filtering employed in spatially-multiplexed multiview 3-D displays (see "Fighting Aliasing Without Regularity").

PLANAR MULTIVIEW DISPLAYS

Two-view stereoscopic displays, while relatively simple and inexpensive, have significant disadvantages for displaying high fidelity 3-D. In glasses-based systems, the same two views are presented to the viewer regardless of viewing position, which can lead to mental confusion about the shape of the object. As already mentioned, two-view autostereoscopic systems based on parallax barriers or lenticulars are also vulnerable to apparent depth inversion of the scene.

> **THE FUTURE ENABLED BY THE COMBINATION OF 3-D SIGNAL PROCESSING, DIGITAL CONTENT, AND NEW DISPLAY TECHNOLOGIES IS ONE OF MANY EXCITING POSSIBILITIES.**

The two-view stereoscopic displays of the type described thus far also fail to provide the important depth cue of motion parallax, i.e., a change of viewpoint due to viewer motion. Motion parallax is an important complement to stereopsis, particularly when a scene is relatively static and the viewer is free to inspect it from a variety of positions. Motion parallax can provide meaningful depth information over a larger range of working distances between viewer and object than can stereopsis, and it provides depth cues to a larger portion of the population (including those with deficiencies of binocular vision).

Several display technologies provide motion parallax in addition to stereoscopic depth by displaying a range of viewpoints of a scene to the viewer. We refer to displays of this type as *multiview displays*. This section describes two main displays of this type: active multiview systems, where a viewer's position is used to calculate and present the appropriate images, and passive multiview displays, where several viewpoints of the scene are projected simultaneously into a viewing zone through optical means.

ACTIVE MULTIVIEW 3-D DISPLAYS

In order to display motion parallax, a device must present view information to an observer that is specific to his or her location. *Active* multiview displays provide motion parallax by tracking the viewer's location, generating two appropriate views for the viewer's eyes, and using an optical mechanism such as those we have already described to present those views to the viewer [11]. An active multiview display, at least ideally, presents the same stereo and parallax cues as would a true 3-D object.

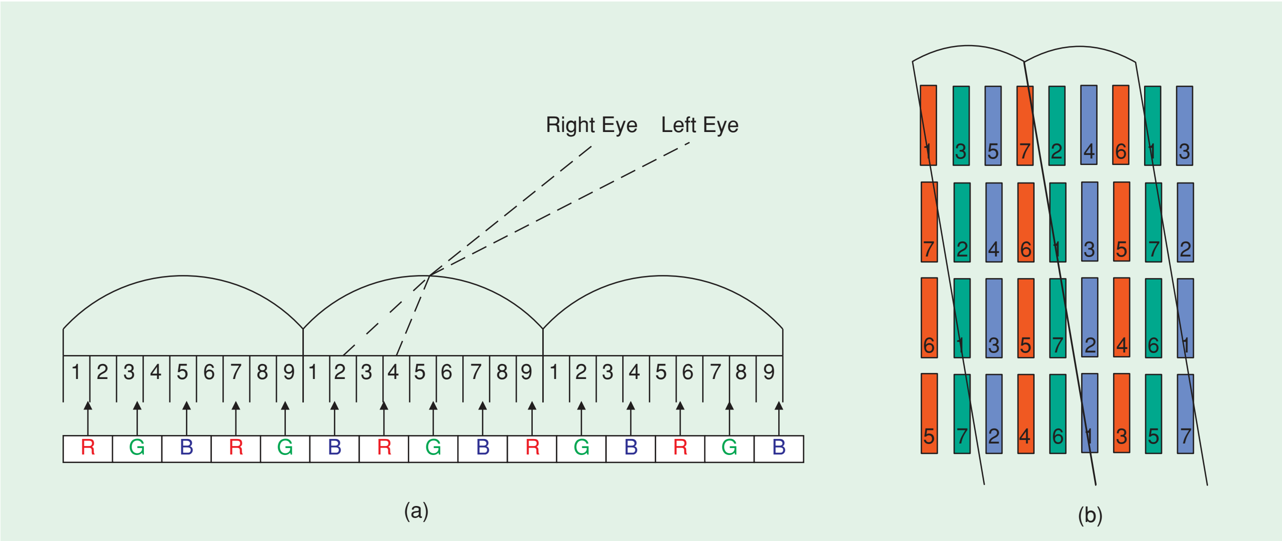
Since view rendering depends on head position, *irregular, temporal* view multiplexing takes place; the degree of irregularity depends on the motion pattern of viewer's head. Since the selection of suitable view for on-screen rendering is akin to sub-sampling of this view's full motion sequence, a temporal anti-alias filter should be employed. However, because of movement-induced irregularity, or time-variance, the design of an optimal filter is nontrivial. Moreover, had one been able to find an optimal filter for this case, another significant deficiency would remain; because of the tracking employed, correct view can be only presented to the viewer being tracked. Essentially, active multiview 3-D displays are single-viewer displays.

PASSIVE MULTIVIEW 3-D DISPLAYS

The inability to support multiple viewers is a major weakness of active multiview 3-D displays. An alternative is to display all views simultaneously by means of spatial multiplexing [13]. If each view (light from the corresponding pixels) is directed into a distinct viewing zone in front of the 3-D display, then viewer motion across zones will induce motion parallax. The viewer must, however, rest at a specified distance in front of the screen, as otherwise the effect is largely lost. Since all views are displayed at the same time, and no head tracking is used, multiple viewers can experience 3-D sensation consistent with their respective positions in front of the display. Support for multiple viewers, motion parallax, and elimination of glasses have proved very appealing and a number of parallax-barrier and lenticular multiview 3-D displays have been successfully launched commercially in the last few years [14]–[16]. Multiview displays that require no glasses are often referred to as *automultiscopic* displays.

The same display technologies used to implement autostereoscopic displays (Section II-B), parallax barriers and lenticulars, are the most commonly used methods for implementing passive multiview systems (Figure 4(a)). The linear nature of the slits or lenses in these displays means they are capable of displaying only horizontal parallax, thus providing depth information to a viewer with eyes horizontal and moving from side-to-side. This type of displays is thus known as *horizontal parallax only*, or *HPO*, display. While full-parallax passive multiview displays have been developed, they are currently poorly suited to electronic implementation because of insufficiently high pixel density of plasma and LCD panels.

In early approaches to multiview digital imaging, the slits or lenses remained aligned with pixel columns, thus inducing N-fold horizontal resolution loss for an N-view display (out of N consecutive pixels in one row each belongs to a different view). This resulted in a severe imbalance between vertical and horizontal resolutions perceived. In order to correct this imbalance, it was later proposed to slightly slant the slits or lenses at a carefully selected angle in order to minimize Moiré patterns resulting from interference of regular structure of the slits/lenticules and screen raster [13] (Figure 4(b)). This tilt causes same-view pixels to be distributed more uniformly (both horizontal and vertical sub-sampling), as shown in Figure 5, however the regular nature of sub-sampling induced



[FIG4] Illustration of the principle of lenticular multiview displays: (a) single-row cross-section of a 9-view display; and (b) impact of lenticule slant on the spatial distribution of same-view pixels for a 7-view display.

by nontilted lenses is lost. This poses a significant problem of aliasing due to irregular sub-sampling. How to characterize such aliasing? How to design anti-alias filters? Some solutions recently developed are presented in "Fighting Aliasing Without Regularity."

PROJECTIONAL MULTIVIEW DISPLAYS

An emerging class of multiview displays use multiple projectors or scanning optics, combined with optics that shape individual view zones, to present multiple views autostereoscopically. Examples of this kind of display include numerous time multiplexed designs from Cambridge University [12], [21], a joint Cambridge/MIT work [22], multi-projector displays using lenticular sheets at MERL [23], and multi-projector displays with holographic optical elements from the Hungarian company Holografika [24]. From a signal processing standpoint, these displays share common characteristics with other multiview displays and with holographic stereograms, and are susceptible to the same types of aliasing artifacts (described in a later section).

On the other hand, projectional multiview displays can maintain the full spatial resolution of each view, rather than directly trading off spatial resolution and number of perspectives the way that lenticular and parallax barrier displays are forced to do. Maintaining full spatial resolution and allowing a more flexible choice of number of view zones is useful when developing displays for high quality cinema or 3-D TV (potentially higher costs of the display system notwithstanding).

INTERPERSPECTIVE ALIASING IN MULTIVIEW DISPLAYS

In addition to signal processing issues related to particular display methodologies, all display types that discretize view space are susceptible to additional aliasing artifacts, specifically interperspective aliasing [25]. Interperspective aliasing results when the spatial frequency of a three-dimensional point in a scene being imaged by a multiview display exceeds that display's permitted resolution. This resolution limit depends on the particular display's geometry, but is always depth-dependent and always related to the size and discretization of the view zone. This resolution limit of the display's optics is exactly analogous to depth of field in an image recording camera.

Failure to bandlimit image content below a display's maximum resolution results in artifacts where parts of the image that should appear as a continuous blur are broken up into a series of dots or image fragments. This artifact can either be prevented during image recording or synthesis [22], [25] or corrected using post-capture filtering [26].

HOLOGRAPHIC DISPLAYS (PLANAR)

Holographic displays are a diverse group of imaging technologies that incorporate diffraction as an underlying component of image formation. Holography is fundamentally different from other display types because of its ability to use not just spatially-varying intensity information, but also phase, to reconstruct the intensity and direction of a light field.

Optical holography was invented in 1948 by Dennis Gabor, but only became practical for imaging in the mid-1960's through Leith and Upatniek's use of the laser. A Leith and Upatniek type hologram is both a recording and a display device. A high resolution photographic plate records the interference between light from an object and a mutually-coherent but otherwise informationless reference light source. In signal processing terms, the reference beam acts as a carrier for the scene's spatial information. This interference pattern is recorded continuously across the entire surface of the recording material.

During viewing, a monochromatic illumination beam, typically at the same relative location to the display as was the original reference source to the recording material, acts to heterodyne the image information down from high frequency spatial information back into an image. Remarkably, the light emerging from the hologram closely approximates the wavefront of the original scene in phase, direction, and intensity. Since this process occurs seamlessly and phase-coherently across the recording material, a hologram becomes a highly accurate three-dimensional window into the original scene. Considered another way, a hologram is an extremely specialized lens that focuses light from the illumination source into the three-dimensional shape of the original scene.

This process of phase manipulation using high-resolution spatial information and diffraction allows holography to exceed

FIGHTING ALIASING WITHOUT REGULARITY

When rendering image data for a passive multiview display, spatial view sub-sampling and multiplexing are required. This is illustrated in Figure 5. Note that red sub-pixels, activated by single view, are not regularly spaced; their locations cannot be described by a 2-D lattice [17] (similarly, for green and blue sub-pixels). Thus, it is not obvious how to quantify the aliasing resulting from such an irregular sub-sampling. Two approaches have been developed to date.

In the first approach [18], the irregular sub-sampling layout $\mathcal{V} \subset R^2$ (red circles in Figure 5) is *approximated in spatial domain* by sub-sampling on a lattice or union of lattices [17], denoted $\Psi \subset R^2$. Clearly, $\mathcal{V} \subset \Gamma \subset R^2$, where Γ is the orthonormal lattice of the screen raster (dots), with sampling matrix $V_\Gamma = \text{diag}(1, 1)$.

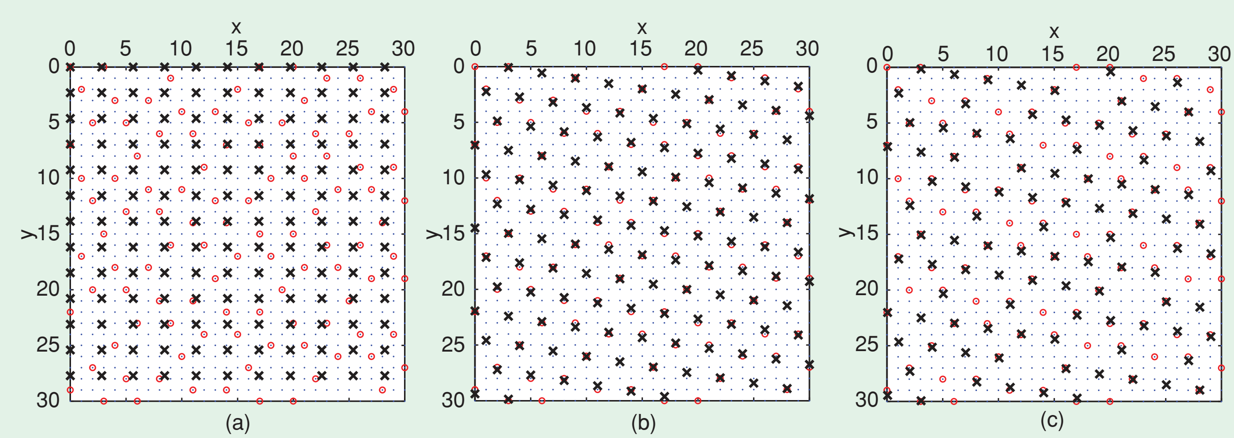
The goal now is to find Ψ such that, in some sense, it best approximates the set $\mathcal{V}$. One possibility is to minimize distance between sets Ψ and $\mathcal{V}$. Let $d(x, \mathcal{A})$ be the distance from point x to a discrete set of points $\mathcal{A}$, i.e., $d(x, \mathcal{A}) = \min_{y \in \mathcal{A}} |x - y|$. Then, a mutual distance between two point sets Ψ and $\mathcal{V}$ can be minimized:

$$\min_{\eta} \beta_{\mathcal{V}} \sum_{x \in \mathcal{V}} d(x, \Psi) + \beta_{\Psi} \sum_{x \in \Psi} d(x, \mathcal{V}) \tag{5}$$

where η is a vector of parameters describing the sampling structure Ψ. For a 2-D lattice, η is a 3-vector (upper-triangular sampling matrix suffices), and thus minimization (5) can be accomplished by hierarchical exhaustive search over a discrete state space. As the weights $\beta_{\mathcal{V}}$ and β_{Ψ} are adjusted, different solutions result. However, if $\beta_{\Psi} = 0$, a very dense (in the limit, infinitely dense) Ψ would result, while for $\beta_{\mathcal{V}} = 0$, a single-point $\Psi \subset \mathcal{V}$ would be found optimal. Instead of a combination of both distances one could use either of them under constrained minimization (e.g., constraint on the density of Ψ).

Applied to various single-view layouts $\mathcal{V}$, the above method has been shown effective in identifying various regular approximations, from orthogonal-lattice approximations (quite inaccurate), through nonorthogonal-lattice approximations (more accurate), to union-of-cosets approximations (most accurate). An example is shown in Figure 5; note a progressively-improved alignment between single-view points from $\mathcal{V}$ (red circles) and model points from Ψ (crosses). Having identified a regular-structure approximation, it is relatively straightforward to identify the passband of suitable low-pass anti-alias pre-filters [17].

In an alternative approach, the irregularity is *approximated in frequency domain* [19]. The main idea is based on the observation that in 1-D a sequence of unit impulses (discrete Kronecker deltas) $g[n]$ has the discrete-time Fourier transform (DTFT) in the form of a periodic impulse train (sum of delayed Dirac delta functions). Therefore, the sub-sampling of a signal $x[n]$, by multiplying it with the sequence $g[n]$, results in the convolution of their DTFTs: $\mathcal{F}\{x[n]\} * \mathcal{F}\{g[n]\}$. Clearly, the spectrum $\mathcal{F}\{x[n]\}$ is going to be replicated at locations of the periodic impulse train $\mathcal{F}\{g[n]\}$. A similar relationship holds in 2-D with respect to bi-sequences and 2-D periodic impulse trains.



[FIG5] Approximation of single-view sub-pixels using: (a) orthogonal lattice; (b) nonorthogonal lattice; and (c) union of 21 cosets. Dots denote the orthonormal RGB screen raster, red circles (∘) denote red sub-pixels of this raster activated when rendering one view on a typical spatially-multiplexed multiview screen, while crosses (×) show model (lattice) locations.

the information storage capacity of all other display technologies. In its purest form, the hologram is free of the limitations of spatial multiplexing and diffraction due to aperturing that can plague other high-resolution multiview display technologies. This property means that holography can actually be used to mimic complex multi-element optical devices, or even to simulate the properties of other 3-D display technologies.

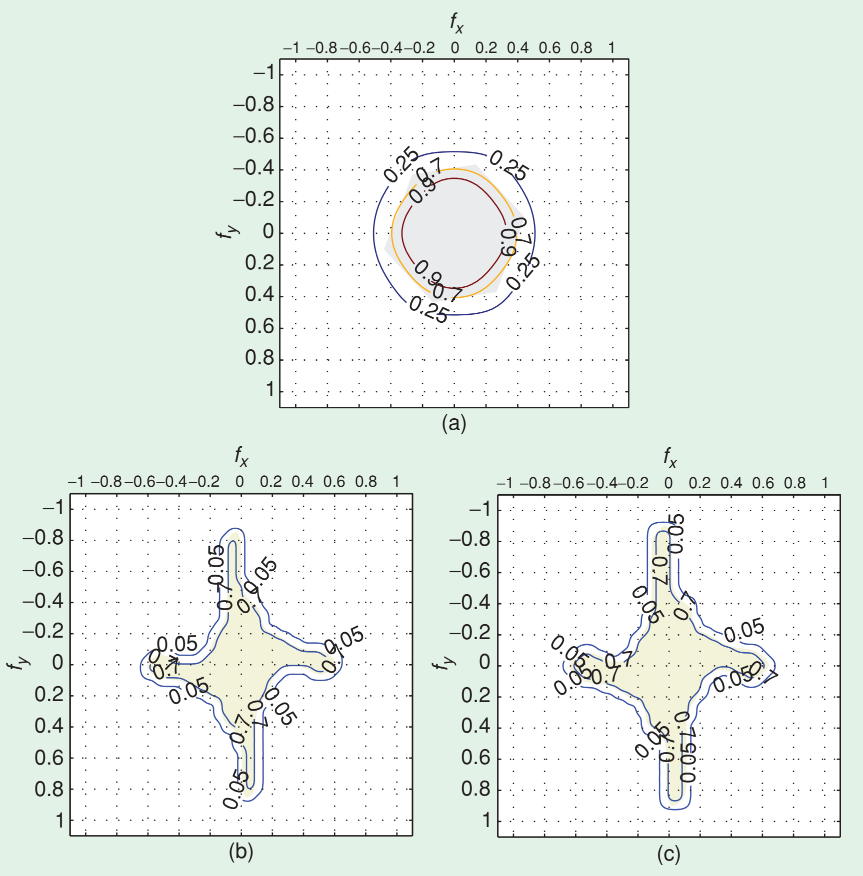
INFORMATION REDUCTION

In reality, the hologram's ability to record extremely large amounts of data is as much its weakness as its strength, particularly when we consider holography as a display for digital or dynamic content. It would be impossible to replicate the detail of a high-quality optically-exposed hologram using a digital recording process. Since this level of image fidelity is simply

The above relationships also hold if $g[n, m]$ is a bi-sequence of *irregularly-spaced* unit impulses, i.e., defined on $\mathcal{V}$. Then, $\mathcal{F}\{g[n, m]\}$ is a 2-D train of Dirac delta functions defined on a reciprocal lattice Λ^* of the least dense lattice Λ such that $\mathcal{V} \subset \Lambda$. Additionally, impulses in this train are scaled differently at different frequency locations [20]. The consequence of this is that the signal spectrum is replicated with different gains at different frequency locations, and replications with smaller gain factors may have less of an impact on aliasing than those with large gain factors. Although to completely prevent aliasing one could consider the worst case scenario, i.e., limiting the signal spectrum so that the closest off-origin spectral replications do not "leak" into the baseband Voronoi cell, such a design would be too restrictive; usually a controlled degree of aliasing is permitted. Then, however, either the actual signal spectrum or its model is needed. Good results were obtained for separable Markov-1 image model [19]; an anti-alias filter's passband boundary was found by identifying frequencies at which the ratio of baseband spectrum magnitude to that of its closest spectral replication (with suitable gain) is either 1 or exceeds 1 (less aliasing).

Figure 6 shows the desired and computed magnitude responses of anti-alias filters using the spatial- and frequency-domain approximations. Note a rotated hexagonal passband for the multiplexing model based on nonorthogonal lattice, and a diamond-shaped passband, with horizontal and vertical extensions, for the frequency-domain approximation. The filter in Figure 6(c) has a slightly larger passband than the one in Figure 6(b) for it takes into account optical crosstalk present in lenticular displays [19]. Applied in practice, pre-filtering based on specifications from Figure 6(b) or 6(c) results in as effective suppression of aliasing artifacts as filtering using specifications from Figure 6(a) while much better preserving horizontal and vertical detail of 3-D images.

The benefit of approximation in frequency domain is that anti-alias filtering can be adapted to specific spectra, whether of a particular image or of a model. In principle, given sufficient computing power (already available in high-end graphics cards), one can imagine anti-alias filtering adapted to each displayed image.



[FIG6] Desired (shaded) and computed (contours) magnitude response of anti-alias filter for multiplexing model: (a) based on nonorthogonal lattice in spatial domain, and based on irregular structure followed by spectrum modeling (b) disregarding optical crosstalk between views, and (c) accounting for such crosstalk. Both frequency axes are normalized to the Nyquist frequency.

not required for most tasks, information reduction is a challenging but essential element of synthetic holography.

Holographic information reduction is possible using a variety of techniques, many drawn from other nonholographic display technologies. These methods include elimination of vertical parallax, pixelization of the display surface, partitioning the display's viewing region into a relatively small number of discrete view zones, and approximating the wavefront of objects at different depths using piecewise cylindrical representations.

The extent of information reduction depends on viewing conditions, human visual response, and the bandwidth of the recording, transmission, computation, and output stages of the display pipeline. For photographic or synthetic displays recorded on holographic film, a discretization known as a holographic stereogram is commonly used. Holographic stereograms approximate a 3-D scene using projectional views, much like those used in a multiview display. These images are recorded optically into the hologram, with each image being visible in only its own view zone. The holographic interference pattern is optically "computed" through interference just like in a traditional hologram. Highly realistic, full parallax, full color, white light illuminated, wall-sized synthetic images are possible using this technology [27]. This optical printing technique has only been applied to static images.

HOLOGRAPHIC VIDEO

Dynamic holographic images, on the other hand, add multiple additional complexities to holographic image formation. For these displays (known as holographic video or electroholographic displays), interference patterns cannot be created optically but instead must be computed and fed through a phase-modulating optical device such as an acousto-optic modulator (AOM). In general, the fringe pattern produced by the modulator is scanned through a view zone using a polygonal or mirror-based scanner. A variety of display systems and computational techniques [28] have been developed since the earliest prototype in 1989–1990 [29]. Holographic fringe computation is now possible using GPU acceleration on modern graphics display cards [30]. However, image modulation devices remain costly limiting the utility of holographic video systems for practical applications at the present time.

An open signal processing question in the field of electroholography is how to decompose the holographic fringe pattern into component elements that lend themselves to a practical blend of inexpensive synthesis, high compressibility, simplified transmission, effective modulation of the display device, or useful optical properties. Unlike their multiview display counterparts or traditional holographic stereograms, electroholographic displays can use a fringe basis that behaves quite differently from a physical lens or aperture. This flexibility opens up the potential for new optimizations and innovations in the display pipeline.

HOLOGRAPHIC OPTICAL ELEMENTS FOR 3-D DISPLAYS

Holography serves an additional, nonimage forming role in three-dimensional display as an optical element in a multiview display device. Since holograms can be thought of as programmable diffractive optics, holographic lenses or diffusers can be fabricated to meet very specific optical requirements in a very compact element. For example, a holographic optical element (HOE) can be used to restrict a view zone horizontally while diffusing it vertically. Furthermore, multiple optical elements can be written into a single holographic element, producing a composite device that behaves as a superposition of its components.

Such complex holographic optical elements can be created through either optical or computational means and replicated repeatedly to make multiple copies of the device. As diffractive elements, though, HOEs are susceptible to chromatic blur, making them most appropriate for monochromatic optical paths or for diffusers where spectral blur is not critical.

VOLUMETRIC DISPLAYS

While surface displays use directional emission of light to simulate the appearance of a 3-D scene, volumetric displays appear to emit light from an entire volume of space. The intensity of each point in the volume can be modulated by an electrical signal. The appearance of a three-dimensional scene can be recreated by drawing it in 3-D in the display volume.

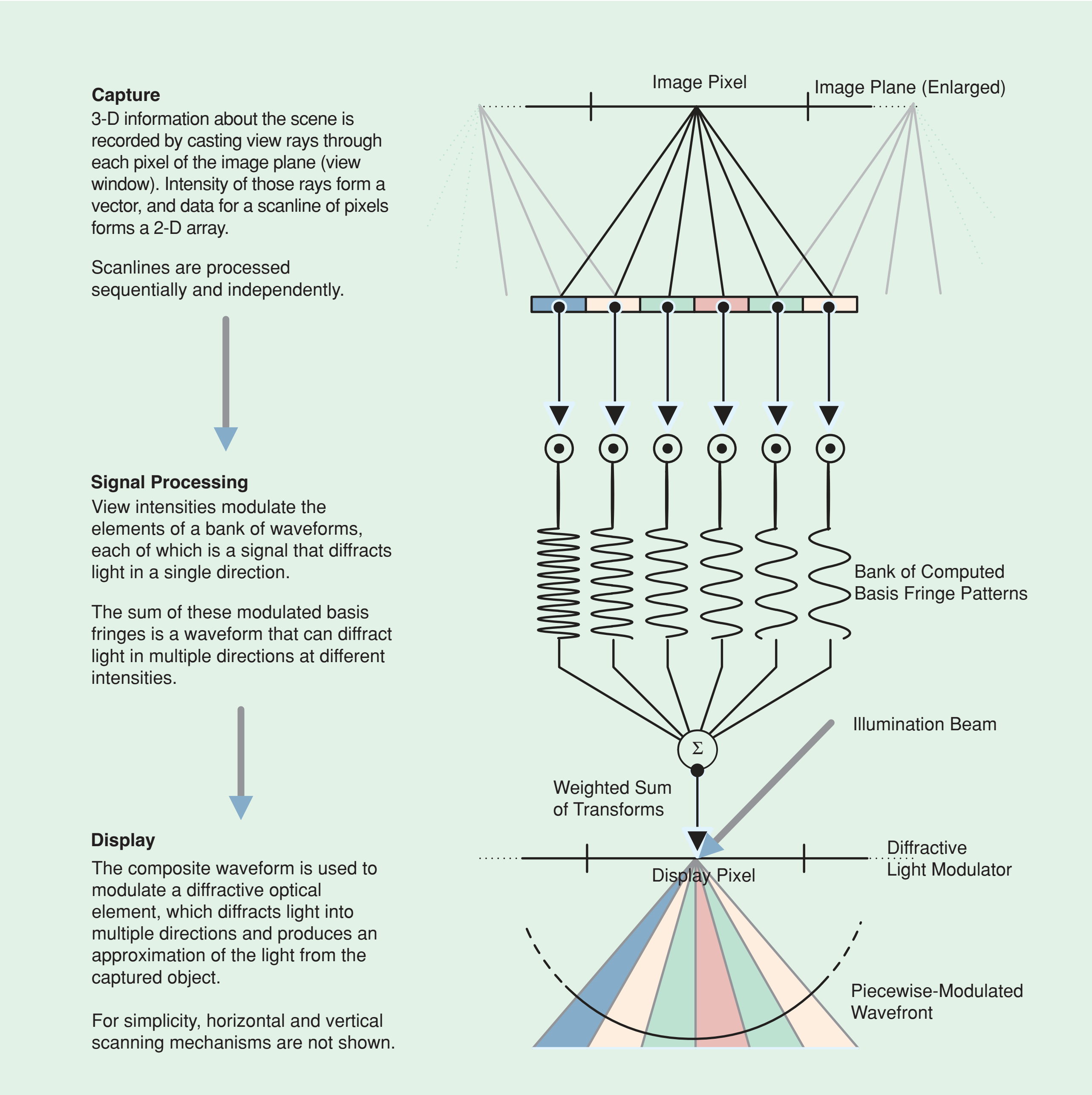
By emitting light from a 3-D location in the volume that directly corresponds to a point in the scene, volumetric displays accurately reproduce the wavefront curvature of light coming from points on the object (from a geometric optics perspective). This property implies that volumetric displays provide visual accommodation (focus) cues to depth, since the eye can focus on the volumetric image of the scene in the same way it would focus on the real scene. Most volumetric displays also reproduce vertical parallax, which is particularly useful for devices that allow scenes to be inspected from a wide range of viewer positions.

Typical volumetric display systems produce image points that have only a single addressable intensity value: each point has an appearance independent of viewer location. On the positive side, this property avoids the interperspective aliasing artifacts of multiview systems since the view zone is not discretized. On the other hand, displays of this type do not maintain accurate occlusion cues for all viewers. Since occlusion is commonly considered the strongest cue to depth in many scenes, the inability to represent occlusion limits the kind of scene in which volumetric displays excel. (Recent work has demonstrated that view-dependence and occlusion can be obtained in volumetric displays; we will describe this technology near the end of this section.)

SWEPT-VOLUME DISPLAYS

Because of the inherent complexity of directly addressing all the points in a three-dimensional volume simultaneously, volumetric displays all employ scanning in some way. Scanning can be accomplished using several different methods. Mechanical volume scanning methods use a passive projection screen or active display element that moves through all parts of the display volume during a single display cycle: these displays are sometimes referred to as *swept-volume* displays. These systems differ in their screen geometry, the shape of their working volume, the order in which this volume is scanned, and the optical configuration of the projector and screen. For example, Perspecta from Actuality Systems [31] uses a flat, double-sided screen to sweep out a hemispherical image volume. The projected image rotates on the screen as the screen sweeps out the image volume. This rotation leads to a relatively complex rasterization ordering of the volumetric image data compared to traditional raster-scanned 2-D displays.

Perspecta also uses a display element very common in contemporary volumetric displays: a DMD (digital micromirror device) image subsystem. DMD elements are extremely useful because of their ability to display image data at extremely high refresh rates (approximately 15 kHz). The binary two-dimensional image information on the DMD can be multiplexed to produce grayscale (through pulse-width modulation), color (using a color wheel), and radial resolution (by varying the ratio of scanning speed to image update). DMD subsystems are available as off-the-shelf devices, replacing what would otherwise be an expensive custom electro-optical display component.



[FIG7] Conceptual block diagram showing how a fringe pattern can be computed from view information and displayed in a holographic video system. The image or view plane is decomposed into pixels through which viewing rays are cast; this view information can be either captured using multiple cameras or computed from a synthetic scene. This view data modulates a set of basis fringe patterns, which are summed to form the input signal for a diffractive light modulator in the display device.

STATIC VOLUME DISPLAYS

Another class of volumetric displays uses electronically addressable elements to scan out the image volume, sometimes known as *static volume displays*. The DepthCube from LightSpace Technologies [32] is an example of such a device. In this display, a DMD imaging system projects a planar image through a stack of electrically-addressable LCD shutters (twenty panels in LightSpace's current version). At any one instant in the display scanning cycle, one of the shutters is active. In this state, the shutter scatters light, acting as a projection screen for the image behind it. Each shutter in the stack is sequentially activated in synchrony with the display of a new projected image, sweeping out an entire image volume.

With twenty depth planes, the depth sampling of the DepthCube is relatively coarse, and gaps between the depth planes are possible to observe. To minimize this problem, LightSpace uses an interpolation method that blurs computationally the gap between adjacent slices, reducing the perceived discontinuity.

OTHER SCANNING METHODS

We mention briefly two other types of volumetric displays that have been developed but to date have not achieved significant commercial success. The varifocal mirror using a moving flexible membrane, acting as a focusing mirror, to opto-mechanically scan an image of a planar display into a viewing zone [33], [34]. The second display type fills the display volume with an optical medium that acts nonlinearly to optical radiation. When two light beams, each with a wavelength invisible to the human eye, cross in the medium, visible light is emitted; the entire image volume can thus be addressed through optical addressing [35], [36].

OCCLUSION AND MID-AIR PROJECTION

Two recent development directions have changed widely spread assumptions about practical volumetric displays. The first assumption was that a volumetric display cannot display view-dependent effects or occlusion. Cossairt et al. [37] modified an existing Actuality Systems Perspecta system with a screen with limited horizontal diffusion. This change permitted view dependence to be displayed. In many ways this display is a hybrid between volumetric and multiview displays, possessing some characteristics of each. While the described device was only a prototype, we are aware of additional, currently unpublished work on this class of displays. Numerous signal processing challenges will doubtless emerge as these displays mature.

The second assumption about volumetric displays was that mid-air display of three-dimensional form of nontrivial depth is not practical without the introduction of a nonlinear display medium. Recent work at AIST in Japan [38] has shown that display comprised of a strongly focused, scanned infrared laser beam is capable of drawing simple 3-D scenes comprised of glowing balls of plasma ionized from air. While the limitations of this system are not yet known, its invention does remind us to periodically revisit common assumptions and reassess their validity in the face of changing technology.

SUMMARY AND CONCLUSIONS

A survey paper such as this one can only give the reader a first glimpse at rich and varied field of 3-D displays. Most of the 3-D technologies we have described here can be traced back to origins well before the dawn of digital signal processing. Different display types have experienced periodic popularity, and some 3-D displays have maintained commercial success in a few specialized areas. Beyond these niches and fads, though, 3-D has never before threatened the dominance of "flat" imaging.

Today, the digital content revolution has changed everything. The CCD has replaced film. The desktop computer has become the director's editing room. Graphics technology has exceeded the traditional animator's wildest dreams. And digital imaging is a component of almost every desktop, every conference room, and every cinema. The digital bit stream is one of the most important foundations for the re-emergence of 3-D, and 3-D specific signal processing is the essential component that allows us to take advantage of it.

The 3-D displays need extra image information to provide a sense of dimensionality, and that data can be processed like any other signal. But it is the 3-D specific algorithm that can, for example, compress multiple views into one by taking advantage of inter-perspective coherence. 3-D aware image processing can compensate for display limitations by minimizing or removing crosstalk between channels. Multiple images can be merged for autostereoscopic displays, with minimal artifacts, by a specialized embedded mosaicing engine. Aliasing artifacts can be filtered out of multiview camera data, again using algorithms tailored to 3-D data.

New signal processing algorithms, new 3-D display devices, and new ways to create 3-D content, are, together, rapidly curing 3-D's longtime ills and ushering in a new era of ubiquitous dimensional imaging. We believe that the early successes we are seeing today, in part because of the kinds of signal processing techniques we have described, will only expand as public interest grows, as commercial successes are achieved, as new algorithms and devices are developed, and as computation becomes ever cheaper.

The future enabled by the combination of 3-D signal processing, digital content, and new display technologies is one of many exciting possibilities. Even today, a volumetric display of cardiac images can allow a team of doctors to treat a beating heart, and audiences in a packed theater can all reach out at once to try to touch synthetic snow in a completely digital 3-D movie. We can only imagine, or perhaps invent, the kinds of applications that will emerge when complete 3-D display systems become widely available.

ACKNOWLEDGMENTS

The authors would like to thank Prof. E. Dubois of the University of Ottawa, Canada for providing Figure 1(a) and valuable comments on early drafts of this paper. Janusz Konrad would like to acknowledge the support of NSF under Grant ECS-0219224, NIH under Grant 1R21HD050655-01, and Boston University under SPRInG Award. Michael Halle would like to acknowledge the support of the NIH NCRR Grant P41-RR13218 and the NIH Roadmap Initiative Grant U54-EB005419.

AUTHORS

Janusz Konrad (jkonrad@bu.edu) received the M.Eng. degree from the Technical University of Szczecin, Poland in 1980, and the Ph.D. degree from McGill University, Montreal, Canada, in 1989. Since 2000 he has been with Boston University. He is an associate technical editor for *IEEE Communications Magazine* and associate editor for the *EURASIP International Journal on Image and Video Processing*. He was an associate editor for *IEEE Transactions on Image Processing* and *IEEE Signal Processing Letters*, a member of the IMDSP Technical Committee of the IEEE Signal Processing Society, as well as the Technical Program cochair of ICIP-2000 and Tutorials cochair of ICASSP 2004. He is a corecipient of the 2001 Signal Processing Magazine and the 2004–2005 EURASIP Image Communications Best Paper awards. He is a Senior Member of the IEEE.

Michael Halle (mhalle@bwh.harvard.edu) is instructor of radiology at Harvard Medical School and Director of Technology Development at the Surgical Planning Lab at Brigham and Women's Hospital. He received his Bachelor's degree in Computer Science from MIT in 1988. During his graduate studies at the MIT Media Lab, he developed computer graphics and image processing techniques that enabled some of the most advanced holographic displays of the time. He received his Master's degree from MIT in 1991 and his Ph.D. in 1997. His research interests include 3-D display, software architecture, and medical visualization. He is also an avid 3-D and "flat" photographer.

REFERENCES

[1] T. Okoshi, *Three Dimensional Imaging Techniques*. Academic Press, 1976.

[2] S.A. Benton, Ed., *Selected Papers on Three-Dimensional Displays*. Bellingham, WA: SPIE Optical Engineering Press, 2001.

[3] N.J. Wade, *Brewster and Wheatstone on Vision*. London and New York: Academic Press, 1983.

[4] E. Dubois, "A projection method to generate anaglyph stereo images," in *Proc. IEEE Int. Conf. Acoustics Speech Signal Processing*, vol. 3, May 2001, pp. 1661–1664.

[5] W. Sanders and D. McAllister, "Producing anaglyphs from synthetic images," in *Proc. SPIE Stereoscopic Displays and Virtual Reality Systems*, Jan. 2003, pp. 348–358.

[6] H. Jorke and M. Fritz, "Infitec—A new stereoscopic visualization tool by wavelength multiplexing imaging," in *Proc. Electronic Displays*, Sept. 2003.

[7] P. Bos, "Time sequential stereoscopic displays: The contribution of phosphor persistence to the "ghost" image intensity," in *Proc. ITEC'91 Annual Conf., Three-Dimensional Image Tech.*, H. Kusaka, Ed., July 1991, pp. 603–606.

[8] J. Lipscomb and W. Wooten, "Reducing crosstalk between stereoscopic views," in *Proc. SPIE Stereoscopic Displays and Virtual Reality Systems*, vol. 2177, Feb. 1994, pp. 92–96.

[9] J. Konrad, B. Lacotte, and E. Dubois, "Cancellation of image crosstalk in time-sequential displays of stereoscopic video," *IEEE Trans. Image Process.*, vol. 9, no. 5, pp. 897–908, May 2000.

[10] J. Fergason, S. Robinson, C. McLaughlin, B. Brown, A. Abileah, T. Baker, and P. Green, "An innovative beamsplitter-based stereoscopic/3D display design," in *Proc. SPIE Stereoscopic Displays and Virtual Reality Systems*, vol. 5664, pp. 488–494, Mar. 2005.

[11] I. Sexton and P. Surman, "Stereoscopic and autostereoscopic display systems," *IEEE Signal Process*. Mag., vol. 16, no. 3, pp. 85–99, 1999.

[12] N. Dodgson, "Autostereoscopic 3-D displays," *IEEE Computer*, vol. 38, no. 8, pp. 31–36, Aug. 2005.

[13] C. Van Berkel, D.W. Parker, and A.R. Franklin, "Multiview 3D-LCD," in *Proc. SPIE Stereoscopic Displays and Virtual Reality Systems*, vol. 2653, pp. 32–39, 1996.

[14] L. Lipton and M. Feldman, "A new stereoscopic display technology: The SynthaGram," in *Proc. SPIE Stereoscopic Displays and Virtual Reality Systems*, vol. 4660, Jan. 2002, pp. 229–235.

[15] A. Schmidt and A. Grasnick, "Multi-viewpoint autostereoscopic displays from 4D-Vision," in *Proc. SPIE Stereoscopic Displays and Virtual Reality Systems*, vol. 4660, pp. 212–221, Jan. 2002.

[16] D. de Boer, M. Hiddink, M. Sluijter, O. Willemsen, and S. de Zwart, "Switchable lenticular based 2D/3D displays," in *Proc. SPIE Stereoscopic Displays and Virtual Reality Systems*, vol. 6490, pp. 64,900R–1–64,900R–8, Jan. 2007.

[17] E. Dubois, "The sampling and reconstruction of time-varying imagery with application in video systems," *Proc. IEEE*, vol. 73, no. 4, pp. 502–*522*, Apr. 1985.

[18] J. Konrad and P. Agniel, "Subsampling models and anti-alias filters for 3-D automultiscopic displays," *IEEE Trans. Image Process*., vol. 15, no. 1, pp. 128–140, Jan. 2006.

[19] A. Jain and J. Konrad, "Crosstalk in automultiscopic 3-D displays: Blessing in disguise?," in *Proc. SPIE Stereoscopic Displays and Virtual Reality Systems*, vol. 6490, pp. 12.1–12.12, Jan. 2007.

[20] A. Jain, "Crosstalk-aware design of anti-alias filters for 3-D automultiscopic displays," Master's thesis, Boston University, Apr. 2006.

[21] A. Travis, "Autosteroscopic 3-D display," *Applied Optics*, vol. 29, no. 29, pp. 4341–4343, 1990.

[22] C.N. Moller and A.R.L. Travis, "Correcting interperspective aliasing in autostereoscopic displays," *IEEE Trans. on Visualization and Computer Graphics,* vol. 11, no. 2, pp. 228–236, 2005.

[23] W. Matusik and H. Pfister, "3D TV: A scalable system for real-time acquisition, transmission, and autostereoscopic display of dynamic scenes," in *Proc. ACM SIGGRAPH'04*, New York, NY, USA: ACM Press, pp. 814–824, 2004.

[24] T. Balogh, Z. Dobranyi, T. Forgacs, A. Molnar, L. Szloboda, E. Gobbetti, F. Marton, F. Bettio, G. Pintore, G. Zanetti, E. Bouvier, and R. Klein, "An interactive multi-user holographic environment," in *Proc. ACM SIGGRAPH'06 Emerging technologies*. New York, NY, USA: ACM Press, p. 18, 2006.

[25] M.W. Halle, "Holographic stereograms as discrete imaging systems," in *Proc. SPIE Practical Holography*, vol. 2176, pp. 73–84, May 1994.

[26] M. Zwicker, W. Matusik, F. Durand, H. Pfister, and C. Forlines, "Antialiasing for automultiscopic {3D} displays," in *Proc. ACM SIGGRAPH '06 Sketches*. New York, NY, USA: ACM Press, 2006, p. 107.

[27] M. Klug, "Display applications of large-scale digital holography," in *Proc. SPIE Holography: A Tribute to Yuri Denisyuk and Emmett Leith*, vol. 4737, pp. 142–149, July 2002.

[28] W. Plesniak, M. Halle, J.V. Michael Bove, J. Barabas, and R. Pappu, "Reconfigurable image projection holograms," *Optical Engineering*, vol. 45, no. 11, p. 115801, 2006.

[29] P. St.-Hilaire, S. Benton, M. Lucente, M. Jepsen, J. Kollin, H. Yoshikawa, and J. Underkoffler, "Electronic display system for computational holography," in *Proc. SPIE Practical Holography*, vol. 1212, pp. 174–182, May 1990.

[30] V.M.J. Bove, W.J. Plesniak, T. Quentmeyer, and J. Barabas, "Real-time holographic video images with commodity PC hardware," in *Proc. SPIE Stereoscopic Displays and Virtual Reality Systems*, vol. 5664, Mar. 2005, pp. 255–262.

[31] G.E. Favalora, "Volumetric 3D displays and application infrastructure," *Computer*, vol. 38, no. 8, pp. 37–44, 2005.

[32] A. Sullivan, "DepthCube solid-state 3D volumetric display," in *Proc. SPIE Stereoscopic Displays and Virtual Reality Systems*, vol. 5291, pp. 279–284, May 2004,

[33] J. Muirhead, "Variable focal length mirrors," *Review of Scientific Instruments*, vol. 32, no. 2, pp. 210–211, Feb. 1961.

[34] A. Traub, "Stereoscopic display using rapid varifocal mirror oscillations," *Applied Optics*, vol. 6, no. 6, pp. 1085–1087, June 1967.

[35] J. Lewis, C. Verber, and R. McGhee, "A true three-dimensional display," *IEEE Transaction on Electron Devices*, vol. ED-18, no. 9, pp. 724–732, Sept. 1971.

[36] E. Downing, L. Hesselink, J. Ralston, and R. Macfarlane, "A three-color, solid-state, three-dimensional display," *Science*, vol. 273, no. 5279, pp. 1185–1189, Aug. 1996.

[37] O. Cossairt, J. Napoli, S. Hill, R. Dorval, and G. Favalora, "Occlusion-capable multiview volumetric three-dimensional display," *Applied Optics*, vol. 46, no. 8, pp. 1244–1250, 2007.

[38] H. Kimura, T. Uchiyama, and H. Yoshikawa, "Laser produced 3d display in the air," in *Proc. ACM SIGGRAPH'06 Emerging Technologies*, New York, NY, USA: ACM Press, p. 20, 2006.