

Resolution of Temporal-Multiplexing and Spatial-Multiplexing Stereoscopic Televisions

Joohwan Kim and Martin S. Banks*

Vision Science Program, University of California, Berkeley, CA 94720, U.S.A.

(Received September 28, 2016 : revised December 2, 2016 : accepted January 3, 2017)

Stereoscopic (S3D) displays present different images to the two eyes. Temporal multiplexing and spatial multiplexing are two common techniques for accomplishing this. We compared the effective resolution provided by these two techniques. In a psychophysical experiment, we measured resolution at various viewing distances on a display employing temporal multiplexing, and on another display employing spatial multiplexing. In another experiment, we simulated the two multiplexing techniques on one display and again measured resolution. The results show that temporal multiplexing provides greater effective resolution than spatial multiplexing at short and medium viewing distances, and that the two techniques provide similar resolution at long viewing distance. Importantly, we observed a significant difference in resolution at the viewing distance that is generally recommended for high-definition television.

Keywords : Stereoscopic 3D displays, Spatial multiplexing, Temporal multiplexing, Resolution
OCIS codes : (120.2040) Displays; (330.5510) Psychophysics; (330.1070) Vision - acuity; (110.2960) Image analysis

I. INTRODUCTION

Resolution is an extremely important component of perceived image quality. Viewing distance is, in turn, important for determining resolution. For example, ITU-R Recommendation BT.709 states that “a high-definition system is a system designed to allow viewing at about three times the picture height, such that the system is virtually, or nearly, transparent to the quality of portrayal that would have been perceived in the original scene or performance by a discerning viewer with normal visual acuity” [1]. At a viewing distance of three times screen height (3 H), the pixel density of the high-definition (HD) format is 56.5 pixels/deg. That density corresponds to a pixel size of slightly more than 1 arcmin, which is considered by practitioners to be equivalent to normal (20/20) visual acuity for a healthy eye; specifically, someone with 20/20 acuity can just read letters with a stroke width of 1 arcmin. To display such letters on a digital device, each pixel should be no larger than 1 arcmin. The reasoning, therefore, is that HD format can properly display the smallest letters that people with normal acuity can read, at the recommended viewing distance.

Stereoscopic (S3D) displays have to show the images for both eyes in one stereo frame. Most S3D displays achieve this by multiplexing the two images either temporally or spatially. Temporal multiplexing alternates left- and right-eye images in time, while spatial multiplexing presents the left-eye image on odd pixel rows and the right-eye image on even rows. With temporal multiplexing, all of the image data are shown to the left eye and none to the right eye at one time, and then all of the image data are shown to the right eye and none to the left at another time. With spatial multiplexing, half of the displayed pixels are shown to the left eye and half to the right eye in a given frame. Because only half of the pixels are displayed, numerous investigators and practitioners have suggested that the effective resolution of such displays is reduced relative to temporally multiplexed displays. Others, in contrast, have argued that effective resolution in such displays is not reduced, because the brain fuses the two monocular images into a full-resolution binocular image [2, 3]. To find out which of these two accounts is more valid, we conducted a psychophysical investigation of how the multiplexing technique affects the effective resolution of the display.

*Corresponding author: martybanks@berkeley.edu

Color versions of one or more of the figures in this paper are available online.



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

There are two temporal-multiplexing methods [4]. In simultaneous-capture, alternating-presentation, left- and right-eye image data are captured simultaneously and presented alternately to the two eyes. In alternating-capture, alternating-presentation, left- and right-eye image data are captured and presented alternately to the two eyes. The two techniques only differ when the stimulus changes over time. In this paper we focus on stationary stimuli, so the distinction is not important.

There are three spatial-multiplexing methods [5, 6]. From top to bottom in Fig. 1, they are same-line allocation, alternate-line allocation, and both-line allocation. In each row, the left eye's image data are shown on the left and the right eye's on the right. The displayed stereo image is in the center. Line 1 in the displayed image is visible to the left eye, and is either line 1 or line 2 from the left eye's image data. Line 2 in the displayed image is visible to the right eye, and is either line 1 or line 2 from the right eye's image data.

In same-line allocation, the same rows in the image data are shown to both eyes. Specifically, the odd rows displayed

to the left eye contain data from the odd rows in the left eye's image data, and the even rows displayed to the right eye contain the data from the odd rows in the right eye's image data. The even rows of the image data are not displayed at all.

Alternate-line allocation recruits the rows in alternating fashion from both eyes' image data. The odd rows in the left eye's image data are displayed as odd rows to the left eye, and the even rows in the right eye's image data are displayed as even rows to the right eye. The even rows in the left-eye image data and the odd rows in the right-eye image data are not displayed.

Both-line allocation presents image data from all rows. Pairs of rows in the left-eye image data are combined to be displayed in odd rows on the display, and row pairs in the right-eye image data are combined to be displayed in even rows. In one common implementation of this method, the data are allocated differently in two successive frames. In the first frame, odd rows in the left-eye data are presented to the left eye on odd display rows, and in the second frame even rows in the left-eye data are presented to the left eye on the same odd display rows. The same occurs for the right-eye data and display, but odd image data are first presented to even display rows. The alternating presentation of the pairs of image data rows is meant to happen quickly enough for the data to be temporally averaged by the visual system.

Kim and Banks [7] measured the effective resolution of temporal and spatial multiplexing techniques, and showed that it was higher with temporal multiplexing at viewing distances of 1.5 and 3 times screen height (respectively, 1.5 H and 3 H). Yun, Kwak, and Yang [8] compared grating visibility with both multiplexing techniques at a distance of 3 H, and found that higher spatial frequencies were visible with temporal multiplexing. However, these studies have two limitations. (1) They tested only one spatial-multiplexing method. Park, Kim, and Choi [9] tested same-line and alternate-line allocation, and found that perceived image quality was the same with the two allocation methods. They did not measure effective resolution. (2) They did not test at the currently recommended viewing distance of 3.2 H, where one pixel subtends exactly 1 arcmin at the viewer's eye [10]. They also did not test at the viewing distance that is recommended by prominent TV manufacturers; for example, LG recommends a distance of two times the screen diagonal, which corresponds to 4.1 H.

In the current work, we addressed both of these limitations by testing different allocation methods for spatial multiplexing, and by testing at distances of 3.2 and 4.1 H.

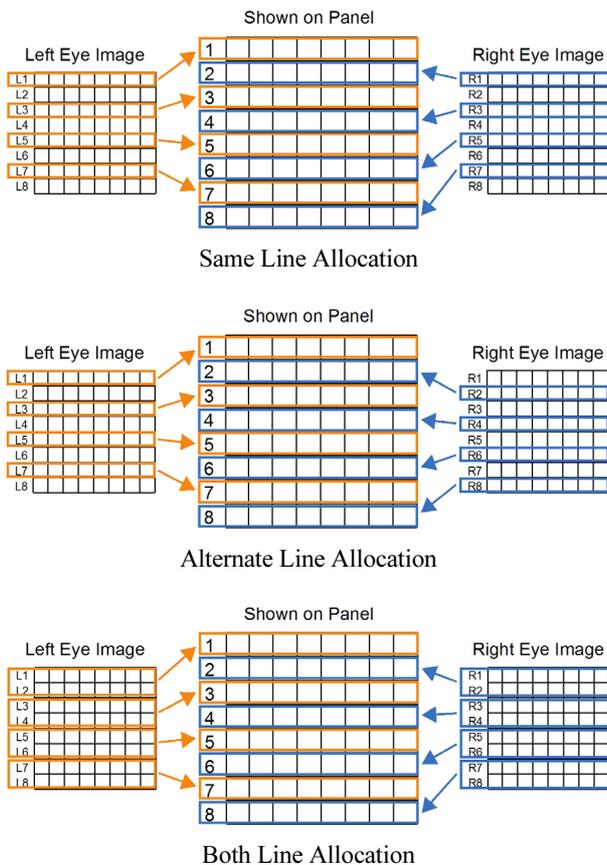


FIG. 1. Three methods of stereo image presentation in spatial multiplexing [5, 6]. Same-line allocation uses the same rows from each eye's image. Alternate-line allocation recruits rows from each eye's image in an alternating fashion. Both-line allocation uses two rows from the left and right eyes' images to generate one row in the stereo image.

II. EXPERIMENT 1: TV COMPARISON

We measured the effective resolutions of two commercial televisions, one using temporal multiplexing and the other using spatial multiplexing. The televisions were set to their

default modes, so they differed somewhat in brightness, contrast, and crosstalk.

Apparatus

The temporal-multiplexing TV was a Samsung LED TV 8000 and the spatial-multiplexing TV was an LG 55LW6500. Both TVs were 55" on the diagonal with 1920×1080 resolution in 2D mode. Both were tested in their default stereo mode. We used the stereo glasses that came with the products. The luminance of the Samsung TV was 133 cd/m^2 without stereo glasses, and 37 cd/m^2 through the glasses. The luminance of the LG TV was 200 cd/m^2 without the glasses, and 87 cd/m^2 through them.

We provided 1920×1080 resolution images for the left and right eye views by enabling NVIDIA 3D Vision on our gpu (GTX 580). Both TVs recognized the protocol and presented the scene without any scaling or clipping. Of course, spatially multiplexing TV cannot provide lossless display of all of the provided pixels. However, that processing happened on the TV side, not on the control pc.

Stimuli

We used a letter acuity test to measure effective resolution. The stimuli were black capital letters from the English alphabet, presented on a white background (Fig. 2). They were created using the design of the letters in a standard clinical eye chart [11]. Letter height was five times greater than letter stroke width, letter width was four times stroke width, and the spacing between letters was twice the letter width. We pre-rendered high-resolution images (400×500) for the 26 letters of the alphabet. During the experiment, the pre-rendered images were resized as desired. We applied cubic interpolation for antialiasing, as described in the *Discussion* section. Three randomly chosen letters were presented on each trial.

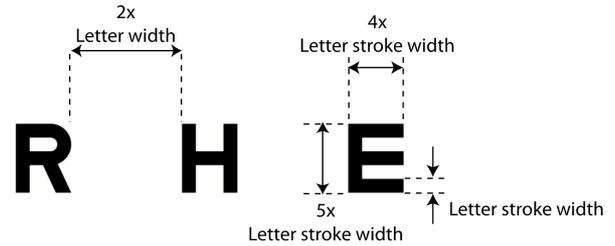


FIG. 2. Stimuli for the visual acuity task. The stimuli followed the design criteria for the most widely used clinical visual acuity test [11]. Letter height was five times letter stroke width. Letter width was four times stroke width. Spacing between letters was two times letter width. Three randomly chosen letters were presented on each trial.

Experimental Conditions

We tested the two multiplexing methods at four viewing distances, yielding eight conditions. The spatial-multiplexing TV used both-line allocation. Each display row on this TV temporally alternated information from odd and even image rows, for a given two-frame sequence. This is equivalent to spatially averaging the two image rows, because the duration of each frame is only 1/120 s, short enough to be fused temporally by the visual system. Figure 3 illustrates this.

We conducted the experiment at four viewing distances: 1.59, 3.18, 4.08, and 6.37 H, which corresponded to 1.09, 2.18, 2.79, and 4.36 m, respectively. Table 1 shows the letter sizes presented at each of the viewing distances. The angular sizes of the letters were roughly the same at all distances, except for the shortest distance, where 3 arcmin was too small to be adequately presented using 2-arcmin pixels.

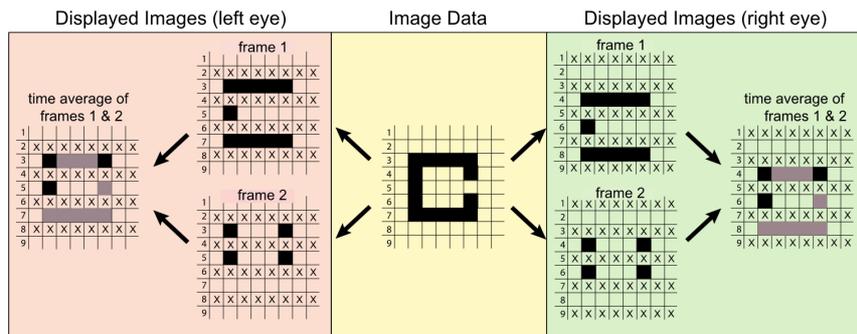


FIG. 3. Schematic of the algorithm used by the spatially multiplexed TV. The middle panel (yellow) shows image data sent to the TV. Black squares represent black pixels in the image data, and bright squares represent white pixels in the image data. The pixel rows are numbered from 1 to 9 so that the reader can keep track of odd and even rows. The left (pink) and right (green) panels show the displayed images presented to the left and right eyes respectively. For the left eye, the first frame presents image data from the odd rows to the odd rows in the display; the second frame presents image data from the even rows to the odd rows in the display. The x's indicate rows that are not seen by the left eye, due to the polarization of the eyewear. The two frames are temporally averaged by the visual system to create the apparent image, labeled "time average of frames 1 & 2." The situation is the same for the right eye, except that the image data are delivered to even rows. Note that the images displayed to the left and right eyes are identical, except that the right-eye image is one pixel row lower on the display screen.

Table 1. Letter sizes presented at the four viewing distances

Viewing distance (picture heights)	Visual angle of one pixel (arcmin)	Angular size of letters (arcmin)	Sampling number in height (pixels)
1.59	2	6, 8, 10, 12, 16, 20, 26	3, 4, 5, 6, 8, 10, 13
3.18	1	3, 4, 5, 6, 8, 10, 13	3, 4, 5, 6, 8, 10, 13
4.08	0.78	2.9, 4.3, 5, 5.7, 7.9, 10, 12.9	4, 6, 7, 8, 11, 14, 18
6.37	0.5	3, 4, 5, 6, 8, 10, 13	6, 8, 10, 12, 16, 20, 26

Procedure

We divided the experiment into roughly 15-minute sessions for different viewing distances and multiplexing methods. The order of sessions was randomized across subjects. After each session, subjects were free to take a break. The whole experiment took about two hours. A total of 3,018 identification responses were obtained from each subject.

On each trial, three letters were presented for 600 ms and then extinguished. This duration is sufficiently long for visual acuity to be maximized [12]. After the stimulus was extinguished, a uniform white stimulus appeared, and the computer awaited the subject's responses before proceeding to the next trial. The subject identified the letters he or she thought were presented by making three keyboard responses, indicating the letters that appeared from left to right on the screen. The procedure was forced choice, meaning that the subject had to make three responses, whether he or she was certain or not. No feedback as to the correctness of the responses was provided. To make sure that the intended keys were pressed, larger versions of the letters associated with each response were written to the screen; the subject could retype a response if the intended key had not been pressed. When the subject was satisfied with the responses, he or she pressed the space bar to proceed to the next trial.

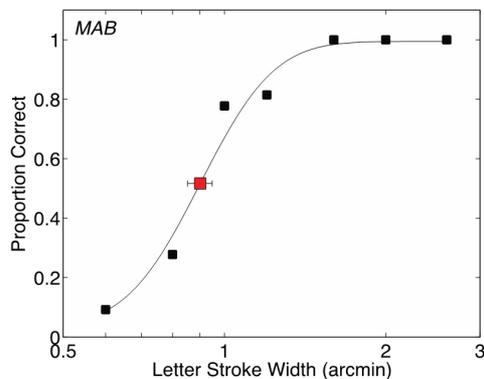


FIG. 4. Psychometric data from one subject in one condition. Proportion of correct letter identification is plotted as a function of letter stroke width. Black squares indicate the data, and the black curve is the Gaussian function that best fits those data. The red square is the estimate of effective resolution. The error bar indicates the 95% confidence interval for the estimate.

Subjects

Six young adult subjects, 24 to 36 years old, participated. All had normal or corrected-to-normal visual acuity and stereoacuity. If they would normally wear corrective lenses, they wore them behind the 3D glasses during the experiment. All but one were naïve to the purposes of the experiment.

Results

We plotted the proportion of correct identification responses as a function of letter size, for each condition. Figure 4 shows data from one subject at one viewing distance. The solid line represents the cumulative Gaussian function that best fit the data:

$$\Psi(x;\theta) = \gamma + (1 - \gamma - \lambda) F(x;\theta) \quad (1)$$

where γ is the guessing rate (1/26), λ is the lapse rate (generally 0), and F is the cumulative Gaussian function [13]. Effective resolution was defined as the letter stroke width at which the proportion of correct responses is 0.50, the steepest part of the function. The red rectangle indicates our estimate of the effective resolution, and the error bar shows the 95% confidence interval of that estimate. Smaller stroke-width values correspond to better resolution.

The data were very similar across subjects, so we averaged them. Figure 5 presents the effective resolution estimates for both multiplexing techniques and all four viewing

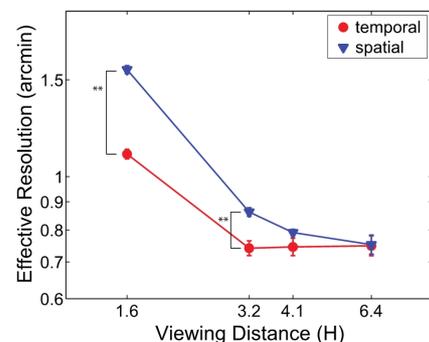


FIG. 5. Effective resolution for temporal and spatial multiplexing. Effective resolution is plotted as a function of viewing distance in units of screen height (H) for the two types of multiplexing. The data have been averaged across subjects. Asterisks indicate statistically significant differences ($p < 0.01$, paired t-test, two-tailed).

distances, averaged across subjects. At 1.6 H and 3.2 H, perceived resolution was significantly better for temporal than for spatial multiplexing ($p < 0.01$, paired t -test, two-tailed). At 4.1 H, temporal multiplexing had slightly better resolution, but the difference was not statistically significant. At 6.4 H, the effective resolutions for the two techniques were essentially identical. One expects effective resolution to be limited by the display at shorter viewing distances, where the pixels are large enough to be resolved by the visual system. At longer viewing distances, however, we expect resolution to be limited by the visual system. The data are consistent with these expectations: Resolution improved with increasing distance from 1.6 to 3.2 H, where one expects performance to be display limited, then leveled off at distances greater than 4.1 H, where one expects performance to be eye limited.

III. EXPERIMENT 2: CRT SIMULATION

The images presented on the two televisions in Experiment 1 differed somewhat in brightness, contrast, and crosstalk. We wanted to make sure that those differences did not cause the observed differences in effective resolution. Therefore we conducted a second experiment in which we simulated the two multiplexing techniques on the same display.

Apparatus

The stimuli were presented on one CRT display using a mirror stereoscope (Fig. 6). By using one display, we could make the luminance, contrast, and crosstalk (in this case there was no crosstalk) identical across conditions. The display was a Viewsonic G255f CRT running at 120 Hz. The screen was 40×30 cm²; pixel size was 0.5 arcmin at the optical distance of 192 cm. The left and right halves of the screen displayed the left- and right-eye images respectively. Maximum luminance was 51.7 cd/m² when viewed via the mirrors. Mirror orientations were adjusted to match the optical and vergence distances of the images.

Stimuli

The stimuli were the same as in Experiment 1, apart from the differences in luminance, contrast, and crosstalk.

Experimental Conditions

We presented three simulated viewing distances (1.59,

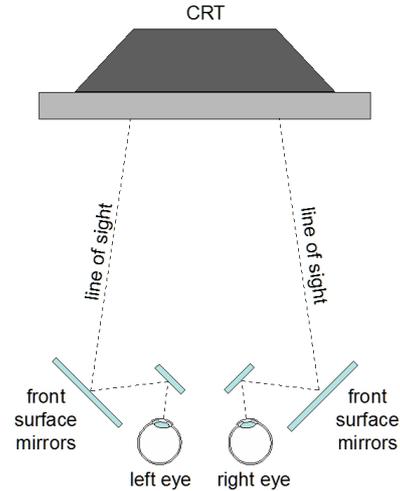


FIG. 6. Apparatus in Experiment 2. A CRT was used to present both eyes' images. The left half of the screen presented the left eye's image and the right half the right eye's image. Four front-surface mirrors created the stereoscopic view at the appropriate distance. The orientations of mirrors were adjusted so that the optical and vergence distances to the images always matched.

3.18, and 6.37 H) by changing the simulated size of pixels (2, 1, and 0.5 arcmin, respectively). We simulated three multiplexing techniques by changing the manner in which images were displayed to the two eyes. The techniques were temporal multiplexing, spatial multiplexing with alternate-line allocation (left eye sees odd rows, right eye sees even), and spatial multiplexing with both-line allocation (each eye sees 60-Hz alternation of even and odd rows in the image data). We excluded spatial multiplexing with same-line allocation, because the resolution will clearly be reduced in that case. The experimental conditions were conducted in double-blind fashion in that neither the experiment nor the subject knew which multiplexing technique was being presented on a given trial. Table 2 summarizes the parameters at each viewing distance.

Procedure

The experiment consisted of one session containing all conditions, presented in random order. It took about two hours to complete.

Subjects

Six subjects, 24 to 39 years old, participated. Three had

Table 2. Letter Sizes in Experiment 2

Viewing distance (screen heights)	Visual angle of simulated pixel (arcmin)	Angular sizes of letters (arcmin)	Sampling number in heights
1.59	2	6, 8, 10, 12, 16, 20, 26	3, 4, 5, 6, 8, 10, 13
3.18	1	3, 4, 5, 6, 8, 10, 13	3, 4, 5, 6, 8, 10, 13
6.37	0.5	3, 4, 5, 6, 8, 10, 13	6, 8, 10, 12, 16, 20, 26

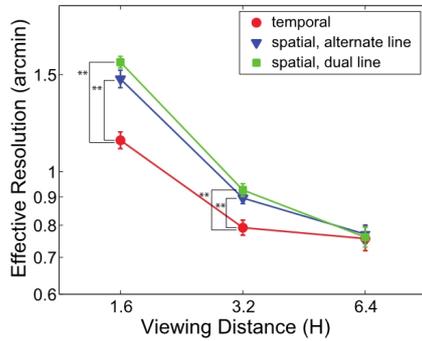


FIG. 7. Effective resolution for simulated temporal and spatial multiplexing. Effective resolution is plotted as a function of viewing distance in screen heights (H) for temporal multiplexing, spatial multiplexing with alternate-line allocation, and spatial multiplexing with both-line allocation. The data have been averaged across subjects. Asterisks indicate statistically significant differences ($p < 0.01$, paired t -test, two-tailed).

participated in Experiment 1. All had normal or corrected-to-normal visual acuity and stereoacuity. If they would normally wear corrective lenses, they wore them during the experiment. All but one were naïve to the experimental purpose.

Results

Figure 7 presents the effective resolutions for different conditions, averaged across subjects. The results were very consistent with those from Experiment 1. At 1.6 H and 3.2 H, temporal multiplexing provided significantly better resolution than either spatial-multiplexing method ($p < 0.01$, paired t -test, two-tailed). At 6.4 H, resolution was not significantly different across the multiplexing methods. Resolution did not differ significantly between the two spatial multiplexing methods. As in Experiment 1, effective resolution improved with increasing viewing distance at the shorter distances, where effective resolution is display-limited. It did not change with distance at the longer distances, where resolution is eye-limited.

IV. DISCUSSION

Summary of Results

It has been argued that spatial multiplexing as implemented in many stereoscopic displays delivers effectively full-resolution binocular images, even though only half of the pixels are presented to each eye at a given time [2, 3]. We found no evidence to support this argument. At viewing distances of 1.6 and 3.2 H, the effective resolution with spatial multiplexing was *not* as high as with temporal multiplexing. This was observed whether the spatial and temporal multiplexing displays were commercially available TVs, or were simulated on a CRT. We also found that

effective resolution was the same for the two multiplexing techniques at longer distances. The results are consistent with the expectation that resolution at shorter viewing distances is limited by the display, and that resolution at longer distances is limited by the eye.

The fact that temporal multiplexing yielded greater resolution than spatial multiplexing at a distance of 3.2 H is significant, because this is the recommended viewing distance for HDTV [10]. However, many viewers tend to sit farther than the recommended distance, so they may experience less reduction in resolution with spatial multiplexing. Indeed, if they sit far enough, they will not experience a reduction at all.

Some previous studies had reported no loss in effective resolution for spatially interlaced TVs compared to temporally interlaced TVs, which is not the same as our finding. We think the most likely reason for the discrepancy is the visual stimuli that were used. Our experiment was designed to measure the threshold value for the two different stereoscopic presentation methods. Note that the subjects were able to score almost 100% when the letter size was larger than the thresholds for both presentation methods. It is possible that the previous studies had used stimuli that were easily recognizable on both presentation methods.

Implications for Ultra-High-Definition (UHD)

The resolution of Full HD was used because that was the prominent resolution format at the time. More recently, TV resolution has been transitioning to UHD. Despite the difference in resolution format, the experimental results and findings are applicable to formats with different resolutions, when converted into angular units at the eye. Visual resolution is determined in angular, not linear, units. This is why, for example, a person's acuity is often stated in MAR (Minimum Angle of Resolution). Indeed, this is why the recommended viewing distance for a resolution format is determined from a calculation using angular units. 3.2 times screen height for HD and 1.6 times screen height for UHD both correspond to 60 pixels per degree [14]. Thus our results yield the same conclusion for HD and UHD: Spatial multiplexing will have lower effective resolution at the recommended viewing distance (3.2 and 1.6 H, respectively), though the difference will diminish at longer viewing distances.

Simulation of Appearance with Temporal and Spatial Multiplexing

The early stages of vision involve low-pass filtering: the eye's optics are aberrated, causing attenuation of high spatial frequencies; the photoreceptors are finite in size, also causing high-frequency attenuation; and photoreceptors are spatially pooled in many higher-order retinal neurons, producing further attenuation of high frequencies. Capturing and displaying a scene also involves low-pass filtering because of pixelation of sensors in capturing devices and pixelation of display panels. Because the visual process occurs after the

display process, the percept derived from a display cannot have greater bandwidth than the direct percept. In other words, a display can at best maintain the bandwidth of direct observation.

Here we examine the pipeline from image capture/creation to image display to viewing in order to determine the expected appearance of images presented on temporally and spatially multiplexed displays for a typical viewer at different distances. Figure 8 shows how the expected appearance will be affected by viewing distance: The upper

and lower panels are for short and long viewing distances, respectively. The bottom graphics in each panel show how the image changes in the frequency domain through the pipeline of displaying and viewing. The process begins with the displayed images on the left. Here the units are cycles per distance on the display plane, so the amplitude spectra are identical because, whether viewed at short or long distance, they are the same physical size. The side lobes are aliases due to finite pixel size. At the eye, we express spatial frequency in angular units of cycles/degree

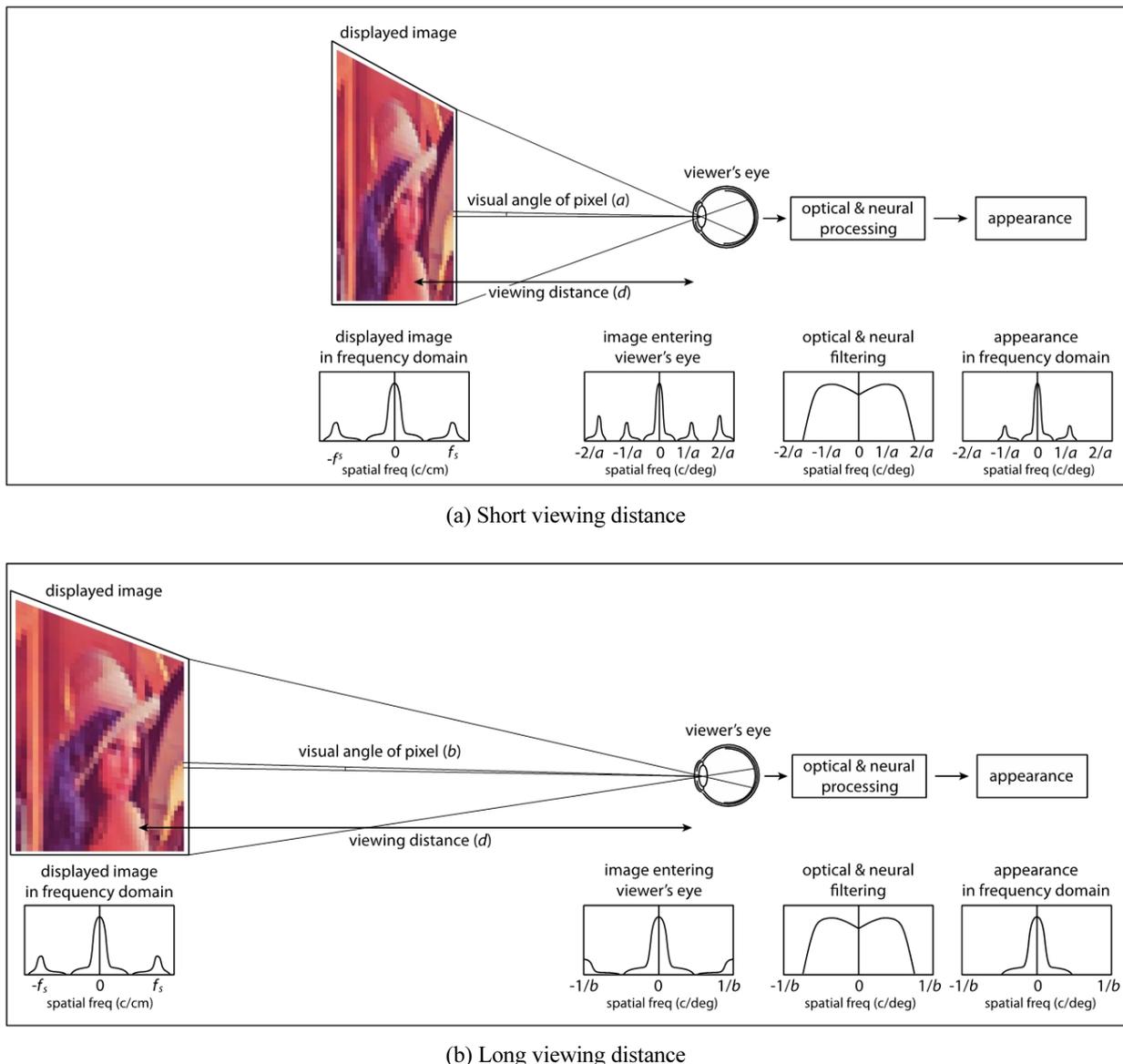


FIG. 8. Visual processing of a scene presented on a display, at short viewing distance (upper panel) and long viewing distance (lower panel). The displayed images are identical in the two cases. The central lobe in the leftmost plot at the bottom of each panel is the signal from the original scene. The side lobes are aliases, due to sampling and displaying. From the viewer's eye, the image subtends a larger angle at short viewing distance and a smaller angle at long distance. As a result, the amplitude spectrum is respectively narrower and wider at those distances (second plots from left). Then the images undergo optical and neural filtering, which we represent by multiplication with the CSF (third plots from left), yielding the amplitude spectra of the image after early visual processing (fourth plots from left). The aliases are still present at the short viewing distance, but have been filtered out at the long distance.

because the low-pass filtering of early vision is to first approximation constant in those units. In angular units, the spectrum narrows and widens for short and long viewing distances, respectively. The image delivered to the eye then undergoes optical and neural filtering. We simulate this with the contrast sensitivity function (CSF) [15]. Specifically, we multiply the amplitude spectrum of the delivered image by the CSF to obtain the amplitude spectrum of the image after visual processing. The visual system behaves linearly near detection threshold and at high spatial frequencies (greater than 2 cycles/deg) [16], so we do not violate the linearity assumption by using the CSF in this way, because our goal is to determine when fine-detail artifacts will be just visible. In the resulting spectra the aliases remain at the short viewing distance, but are eliminated at the long distance. Thus, in this example, the display resolution is not high enough at short distance to avoid aliasing, but it is high enough at long distance.

With an HD TV, temporal multiplexing delivers 56.5 pixels/deg at a viewing distance of 3 H. Spatial multi-

plexing delivers 56.5 pixels/deg horizontally, but only 29.3 pixels/deg vertically. To determine how appearance is likely to be affected in each of the two multiplexing techniques, we go through the procedure in Fig. 9.

The intensity distribution function of the original image is

$$o(x, y), \quad (2)$$

where x and y are respectively the horizontal and vertical coordinates in the spatial domain. In the frequency domain, it is

$$O(f_x, f_y) = \mathfrak{F}[o(x, y)], \quad (3)$$

where \mathfrak{F} is Fourier transformation and f_x and f_y are the horizontal and vertical frequencies. When we down-sample the original, we apply antialiasing using cubic-convolution interpolation [17]. Its kernel is

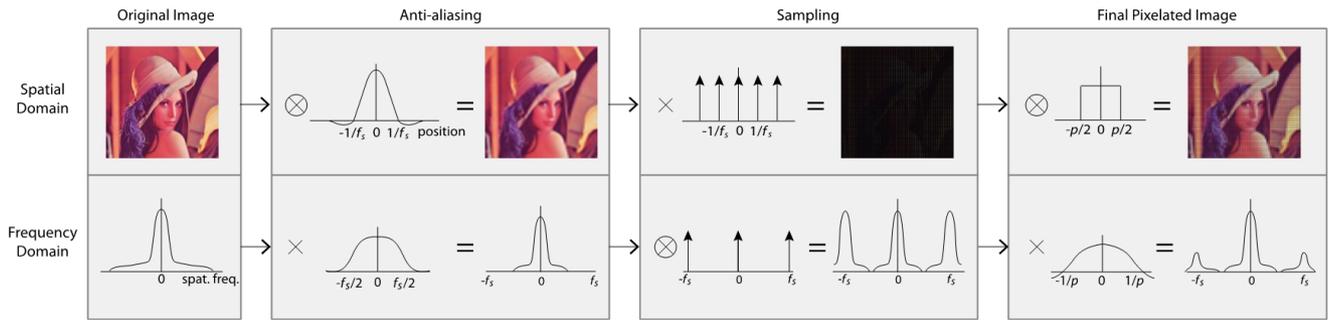


FIG. 9. Sampling and display process. The original image goes through antialiasing and sampling. The antialiasing filter in the second panel is a cubic-convolution interpolation kernel. Sampling is represented in the third panel. In the spatial domain, this is equivalent to multiplying by an impulse-train function with a period of $1/f_s$. In the frequency domain, it is equivalent to convolving with an impulse-train function with period f_s . The sampled information is presented on a display with finite pixel size. In the spatial domain, this is equivalent to convolving the sampled information with a rectangular function whose extent is the same as a pixel on the display. In the frequency domain, it is equivalent to multiplying by the Fourier transform of the rectangular function.

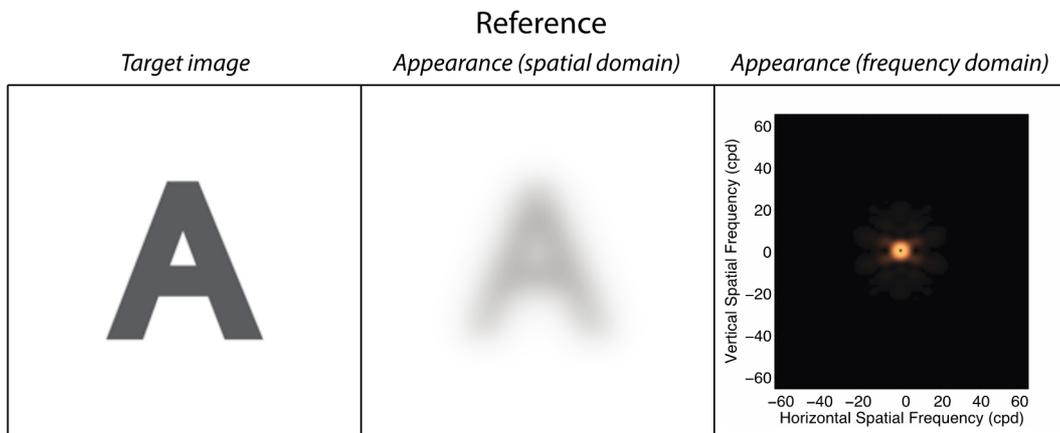


FIG. 10. Simulation of appearance without discrete sampling. The target image, which is the letter 'A', is 10×8 arcmin. The middle panel is the simulated appearance in the spatial domain. The right panel is the appearance in the frequency domain.

$$\begin{aligned}
w(x', y') &= w_x(x') w_y(y') \\
w_x(x') &= \begin{cases} 1.5|x'|^3 - 2.5|x'|^2 + 1 & \text{for } |x'| \leq 1 \\ -0.5|x'|^3 + 2.5|x'|^2 - 4|x'| + 2 & \text{for } 1 < |x'| < 2 \\ 0 & \text{otherwise} \end{cases} \\
w_y(y') &= \begin{cases} 1.5|y'|^3 - 2.5|y'|^2 + 1 & \text{for } |y'| \leq 1 \\ -0.5|y'|^3 + 2.5|y'|^2 - 4|y'| + 2 & \text{for } 1 < |y'| < 2 \\ 0 & \text{otherwise} \end{cases}
\end{aligned} \tag{4}$$

where x' and y' are defined with respect to the sampling period T_s as

$$x' = \frac{x}{T_s}, y' = \frac{y}{T_s}. \tag{5}$$

The intensity distribution of the antialiased image is then

$$a(x, y) = \frac{o(x, y) \otimes w\left(\frac{x}{T_s}, \frac{y}{T_s}\right)}{N_w}, \tag{6}$$

where the normalization factor N_w is

$$N_w = \int_{-2T_s}^{2T_s} \int_{-2T_s}^{2T_s} w\left(\frac{x}{T_s}, \frac{y}{T_s}\right) dx dy. \tag{7}$$

We then sample the antialiased intensity distribution by multiplying the signal and an impulse train:

$$s(i, j) = \sum_{mn} a(iT_s, jT_s) \delta(i - m, j - n), \tag{8}$$

where s is the image data and i and j are the horizontal and vertical indices. The sampled data then gets displayed on a digital device. To represent this, we convolve the image data s with the rectangular function corresponding to the pixel size:

$$d(u, v) = s\left(\frac{u}{T_p}, \frac{v}{T_p}\right) \otimes \text{rect}\left(\frac{u}{L_p}, \frac{v}{L_p}\right) \tag{9}$$

where u and v are the horizontal and vertical angular coordinates on the display, L_p is the pixel size, T_p is the period of the pixel structure, and d is the intensity distribution at the eye. As before, we represent filtering in

early visual stages with the CSF. Specifically, we estimate the visibility of the displayed image by multiplying the spectrum of that image by the CSF:

$$P(f_u, f_v) = D(f_u, f_v) C(f_u, f_v). \tag{11}$$

where D is the Fourier transform of d , C is the CSF from Campbell and Robson [15] and f_u and f_v are horizontal and vertical spatial frequencies. The inverse Fourier transform yields the estimated appearance,

$$p(u, v) = \mathfrak{F}^{-1}[P(f_u, f_v)]. \tag{12}$$

Using this procedure, we now compare temporal and spatial multiplexing at three viewing distances (1.6, 3.2, and 6.4 H, which yield respectively pixel sizes of 0.5, 1, and 2 arcmin). The target image is the letter 'A'; its size is fixed at 10×8 arcmin for all viewing distances. Figure 10 shows how appearance is affected when the letter is directly viewed (*i.e.*, no discrete sampling). The left panel is the original image, the middle panel is the expected appearance in the spatial domain, and the right panel is the appearance in the frequency domain. For the appearance panels to be valid for the reader, the viewing distance must be 7.9 m.

Figure 11 shows the simulation for temporal multiplexing for one eye. We halved the intensity of the target image to account for time multiplexing. From top to bottom, the viewing distances are 1.6, 3.2, and 6.4 H, corresponding to respective pixel sizes of 2, 1, and 0.5 arcmin. The fill factor $Q = (T_p/L_p)^2$ is 0.64 in all cases. As viewing distance increases, sampling is denser and the appearance of the 'A' becomes more veridical. The visibility of the pixel grid is an important consideration. Pixels of 2, 1, and 0.5 arcmin correspond to fundamental frequencies of 30, 60, and 120 cycles/deg respectively. As you can see in the right panels, 30 cycles/deg should be barely visible (four components are present at that frequency), while 60 and 120 cycles/deg should not. Thus, with temporal multiplexing the pixel grid will generally be seen at distances shorter than the recommended distances, but not seen at the recommended and greater distances.

Figure 12 shows the simulation of spatial multiplexing for one eye. From left to right, the columns show the displayed image, its appearance in the spatial domain, and its appearance in the frequency domain. The appearance at viewing distances of 1.6 H and 3.2 H is notably poorer for spatial multiplexing than for temporal, due to the coarser vertical sampling. The pixel rows are more visible with spatial multiplexing, which is apparent in the right panels, where the vertical frequencies due to sampling are visible at the two shorter distances.

Our simulation shows how two widely used techniques for presenting stereoscopic imagery are likely to affect

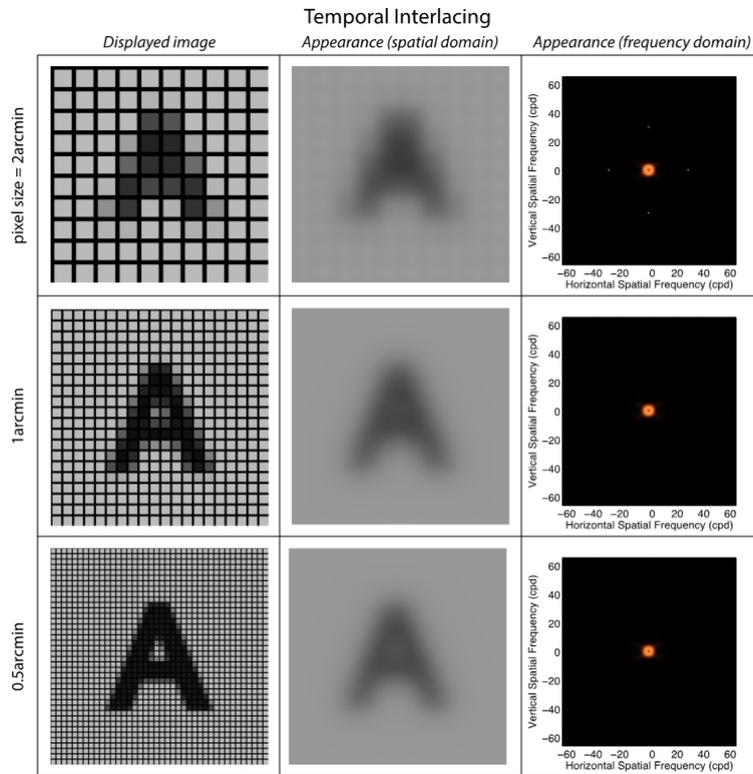


FIG. 11. Simulation of temporal multiplexing. From left to right are the target image presented on the display, its appearance in the spatial domain, and its appearance in the frequency domain. From top to bottom are the results for pixel sizes of 2, 1, and 0.5 arcmin.

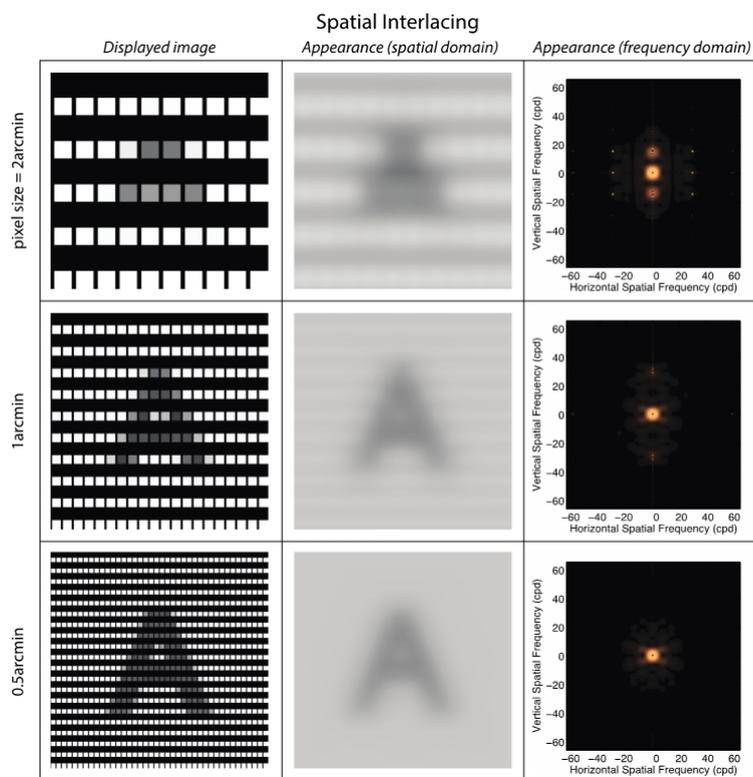


FIG. 12. Simulation of spatial multiplexing. From left to right are the target image presented on the display, its appearance in the spatial domain, and its appearance in the frequency domain. From top to bottom are results for pixel sizes of 2, 1, and 0.5 arcmin.

visual appearance. The simulations were done for one eye, so they do not necessarily inform us about appearance with binocular viewing, but our psychophysical results show quite clearly that effective resolution is lower with spatial multiplexing than with temporal multiplexing at short viewing distances. Those results are actually quite consistent with our simulations and suggest therefore that effective resolution can be well modeled by consideration of the monocular images.

V. CONCLUSION

We compared the effective resolutions of two stereoscopic 3D presentation methods, namely temporal multiplexing and spatial multiplexing. At short and medium viewing distances, temporal multiplexing provided greater effective resolution than spatial multiplexing. At long viewing distance, the two techniques provided similar resolution. The difference in resolution was significant at the recommended viewing distance for HD televisions.

REFERENCES

1. ITU-R Recommendation BT.709-5: "Parameter values for the HDTV standards for production and international programme exchange," International Telecommunication Union, Geneva, Switzerland, 2002.
2. E. F. Kelley, "Resolving resolution," *Information Display* **27**, 18-21 (2011).
3. R. M. Soneira, "3D TV display technology shoot-out," http://www.displaymate.com/3D_TV_ShootOut_1.htm
4. D. M. Hoffman, V. I. Karasev, and M. S. Banks, "Temporal presentation protocols in stereoscopic displays: Flicker visibility, perceived motion, and perceived depth," *Journal of the Society for Information Display* **19**, 271-297 (2011).
5. S. Dawson, "Passive 3D from the beginning," <http://hifi-writer.com/wpblog/?p=3797>
6. S. Dawson, "Active versus passive," *Connected Home Australia*, 46-48 (Sep 2012).
7. J. S. Kim and M. S. Banks, "Effective Spatial Resolution of Temporally and Spatially Interlaced Stereo 3D Televisions," *sid symposium digest of technical papers* **43**, 879-882 (2012).
8. J. D. Yun, Y. Kwak, and S. Yang, "Evaluation of perceptual resolution and crosstalk in stereoscopic displays," *Journal of Display Technology* **9**, 106-111 (2013).
9. M. Park, J. Kim, and H. J. Choi, "Effect of interlacing methods of stereoscopic displays on perceived image quality," *Applied Optics* **53**, 520-527 (2014).
10. ITU-R Recommendation BT.2022: "General viewing conditions for subjective assessment of quality of SDTV and HDTV television pictures on flat panel displays," International Telecommunication Union, Geneva, Switzerland, 2012.
11. I. L. Bailey and J. E. Lovie, "New design principles for visual acuity letter charts," *American Journal of Optometry and Physiological Optics* **53**, 740-745 (1976).
12. W. S. Baron and G. Westheimer, "Visual acuity as a function of exposure duration," *Journal of the Optical Society of America* **63**, 212-219 (1973).
13. I. Fründ, N. V. Haenel, and F. A. Wichmann, "Inference for psychometric functions in the presence of nonstationary behavior," *Journal of Vision* **11**, 11 (2011).
14. ITU-R Recommendation BT.1845-1: "Guidelines on metrics to be used when tailoring television programmes to broadcasting applications at various image quality levels, display sizes and aspect ratios," International Telecommunication Union, Geneva, Switzerland, 2010.
15. F. W. Campbell and J. G. Robson, "Application of Fourier analysis to the visibility of gratings," *The Journal of Physiology* **197**, 551-566 (1968).
16. F. W. Campbell, E. R. Howell, and J. R. Johnstone, "A comparison of threshold and suprathreshold appearance of gratings with components in the low and high spatial frequency range," *The Journal of Physiology* **274**, 193-201 (1978).
17. R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Transactions on Acoustics, Speech and Signal Processing* **29**, 1153-1160 (1981).