Antialiasing of Interlaced Video Animation

John Amanatides and Don P. Mitchell

AT&T Bell Laboratories
Murray Hill, NJ 07974

1. Abstract

The production of computer-generated video presents a number of difficulties not encountered with motion pictures. Interlaced scanning and the color subcarrier of NTSC video are responsible for special problems such as interline flicker, and chroma aliasing. As in motion pictures, temporal aliasing is also an issue. A renderer can sample and filter a moving image in an arbitrary manner and is not constrained to simply imitate the behavior of a television camera. This paper explores several different spatiotemporal antialiasing filters and how they affect the quality of video animation.

CR Categories and Subject Descriptions: I.3.3 [**Computer Graphics**]: Picture/Image Generation

General Terms: Algorithms

Additional Keywords and Phrases: Animation, Antialiasing, Interlacing, Video

2. Introduction

Computer-generated images can be created for a number of different formats, such as still images, motion pictures, or video. It is known today that frames of a motion picture or a video should not necessarily be rendered in the same way as still images, with no consideration of time and motion. This paper focuses on the problem of rendering antialiased video, a problem which has not received much attention in the graphics literature.

In many ways, a motion picture is a simpler representation of moving images than video. A motion picture is made up of a sequence of essentially continuous two-dimensional images (ignoring the effects of film grain). In other words, a motion picture is continuous in the vertical and horizontal dimensions, but discrete in the time dimension. Most of the computer-graphics literature on temporal antialiasing has been concerned with production of motion pictures [Cook84].

Unlike motion pictures, a video signal represents the *scanning* of a moving image. A video signal is essentially continuous in the horizontal dimension, but discrete in the vertical and time dimensions. In addition, this scanning is *interlaced,* meaning that a video signal alternates between scanning the odd numbered lines and the even numbered lines [Tonge84].

In the video domain, both spatial and temporal aliasing can result, as well as other forms of aliasing caused by color encoding. Many of these unwanted aliasing artifacts occur at the source of the video signal, which in most cases is a television camera. Some problems are artifacts of the interlaced TV display.

For computer-generated video, the source is a rendering program which converts a symbolic model of a scene into a digital image. A synthetic video signal could be produced by faithfully simulating a television camera, but rendering software is not constrained by the physical realities that constrain the design of a camera. Given this freedom, it should be possible to synthesize video that is as good or better than the output of a camera.

This work is motivated by our experimental ray tracing program, FX. FX is designed to render images in multiple formats, including video. Moving scenes are represented by the constructive solid geometry (CSG) scheme, augmented with operators for translational and rotational velocity. Internally, samples of a moving image can be made at variable points in (x, y, t). This

ray tracer was written to run on a parallel computer [Potmesil89], and this has helped us to perform computationally expensive experiments in a reasonable amount of time.

3. Theory

A video signal represents the scanning of a moving 2D image at the rate of 30 frames per second (25 in some countries). This scanning is *interlaced*, which means the odd-numbered lines are scanned in the first 1/60th second, and then the even lines are scanned. Figure 1 illustrates this scanning process in a slice perpendicular to the horizontal scan lines. This view in (t, y) coordinates illustrates the discrete sampling nature of the scanning process.
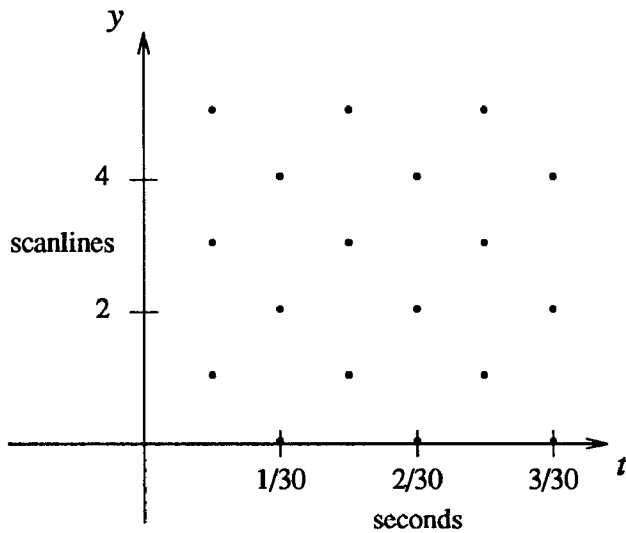

Figure 1. (t, y) Slice Through Interlaced Scans

Even an analog video signal is discrete in the vertical and temporal dimensions, but digital video is also discrete in the horizontal dimension. A frame from a computer graphics video animation invariably starts out as a digital image in a frame buffer. It may continue to be digital throughout the recording and editing process, if the standard 4:2:2 video format is used [CCIR601].

Sampling in the (x,y) dimensions is not interlaced like (t, y)—the sampling pattern is the usual rectangular arrangement of pixels. In our work, we have used the 4:2:2 standard, in which the signal is represented by 720 samples (pixels) in the horizontal dimension, 486 samples (scan lines) in the vertical dimension, and 30 samples (frames) per second in the time dimension.

When a signal is sampled, the resulting interlaced video spectrum consists of replicas of the spectrum of the original signal. Figure 2 shows the spectral consequences of interlaced sampling.
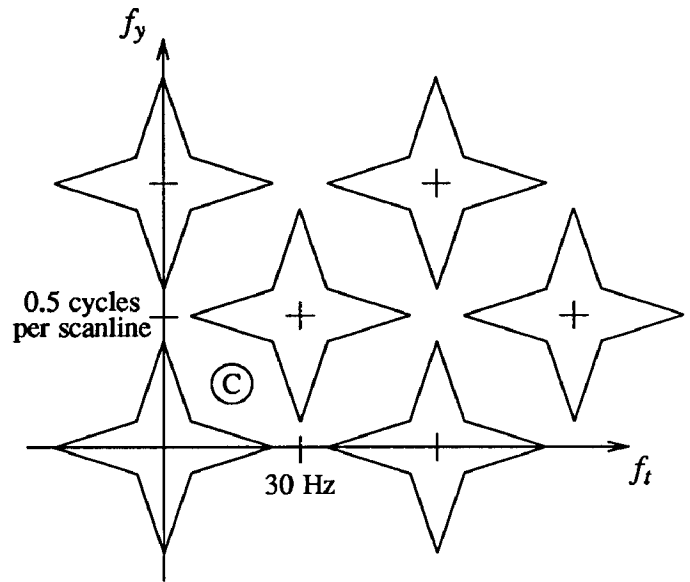

Figure 2. (t, y) Spectrum of Interlaced Digital Video

Figure 3 is another view of the sampled (i.e., digital) video spectrum on the (x,y) directions. Just the replicas centered on the $f_t = 0$ plane are shown:
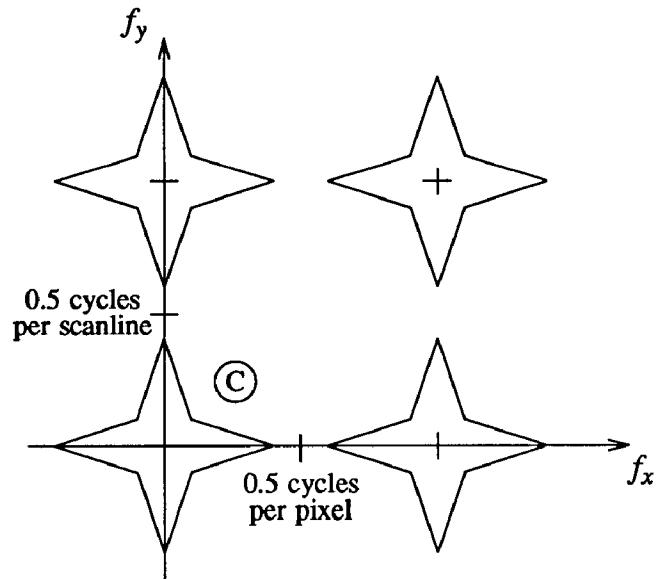

Figure 3 (x, y) Spectrum of Interlaced Digital Video

When the bandwidth of the image spectrum is too wide (in any dimension), the replicas in the sampled spectrum may overlap. This is the origin of aliasing, which can be prevented by prefiltering the signal to limit the bandwidth of the spectrum. Several distinct types of aliasing can occur in the generation of a video signal.

Spatial Aliasing, all too familiar in computer graphics, will occur if the baseband spectrum shown in Figure 3 overlaps with the sideband above it or beside it. Little

will be said in this paper about spatial aliasing, since it is not such a novel problem.

Temporal aliasing will occur if there is overlap (see: Figure 2) with the sideband to the right at (60 Hz, 0 cycles per scanline). This type of problem also appears in motion pictures.

Spatiotemporal aliasing is what we will call the interlace-related problems that result from the sideband in Figure 2 which is to the upper right of the baseband at (30 Hz, 0.5 CPS). In particular, energy at vertical frequences in the baseband near 0.5 CPS will be replicated by the sideband at about 30 Hz. For example, if a pattern is brighter at every other scan line, the display will flicker at 30 Hz.

Chroma aliasing is a problem that can occur in broadcast video. Except in digital or component analog video (used mainly in the editing stage), low-bandwidth color information is encoded into a high-frequency portion of the spectrum. The circled "C" in Figures 2 and 3 show the location and approximate width of this color subcarrier. Energy in the baseband that overlaps this location can cause waves of false color to appear.

4. Temporal Antialiasing

Both video and motion pictures represent movement by discrete samples in time. As with any sampling process, energy above the Nyquist frequency will lead to aliasing. Anyone who has played with a strobe light is familiar with the appearance of temporal aliasing, and the shutter of a movie camera can create similar effects. Moving objects seem jerky, or a spinning wheel may seem to turn backwards. These effects are called "motion judder" or "strobing".

This aliasing can be removed by temporal prefiltering, which is often referred to as *motion blur*. But motion blur is not a panacea. Temporal filters may prevent motion judder, but a moving object may look unnaturally smeared along the direction of motion.

This is a particular problem if viewers track a moving object with their eyes. Motion blur looks right when the viewer is not following the object with his or her gaze. But during visual pursuit, the image of the object is more or less fixed on the retina, and then motion blur looks quite unnatural. Unfortunately, the only way to eliminate these problems would be to greatly increase the frame rate of motion pictures and video (temporal-aliasing effects are visible even at 120 frames per second [Hsu85]). Given the current standardized frame rates, judgements about the degree of motion blur versus judder are generally made by the cinemato-

grapher or videographer; they are not really engineering problems.

A relationship exists between spatial and temporal filtering in the simple case of constant velocity motion (of the image on the viewplane, that is). In this case, when a temporal filter is applied, the appearance of a frame is the same as if a one-dimensional spatial filter were applied to a stationary image along the direction of motion. This principle has been used to implement motion blur via spatial filtering [Potmesil83], but the algorithm is not completely general.

The design of spatial filters involves image-quality trade-offs between aliasing, blurring, and ringing [Mitchell88]. The subjective effects of temporal filtering and sampling is also a complex subject for study [Hsu85]. Temporal aliasing could be removed by strong filtering, but such filters would have to span many frames and would lead to unacceptable blurring or ringing artifacts in moving objects.

Figure 4 illustrates approximately how a CCD television camera filters an image by integrating over a rectangular region of space/time. The sensors in the CCD array integrate over regions one scanline high, and they are ganged together to average pairs of scanlines [Murata83]. This *interlaced box* filter averages over a rectangular region two scanlines high, one pixel wide, and one field-interval (1/60th second) in duration.
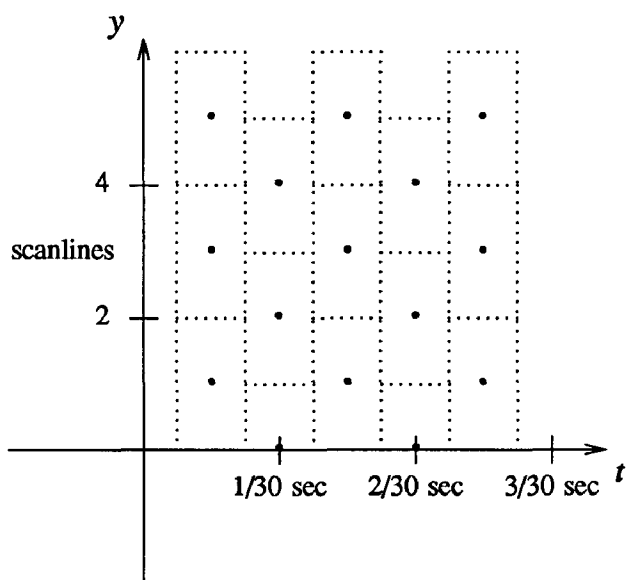

Figure 4. Interlaced Spatiotemporal Box Filter

In Figure 5, two images of the Fresnel zone plate are shown: on the left produced by an actual CCD camera, and on the right synthesized by FX using the interlaced box filter. Synthetic images were filtered by stochastic

sampling with 64 samples per pixel.

Figure 6 shows a frame from a test animation, six humanoid figures revolving on a carousel. At 10 r.p.m., the figures move fairly rapidly in the foreground of the scene, but they can be tracked by the eye.

It is interesting to consider other possible spatiotemporal filters besides the interlaced box. Restricting our attention to volume-integration filters (unweighted average over some volume of space/time), there are still a number of volume shapes that could be used.

One alternative would be like Figure 4, but with the box filters rotated 90 degrees. This filter would average over only one scanline of height and a full frame duration (1/30th second). Unfortunately, this filter yields far too much motion blur.

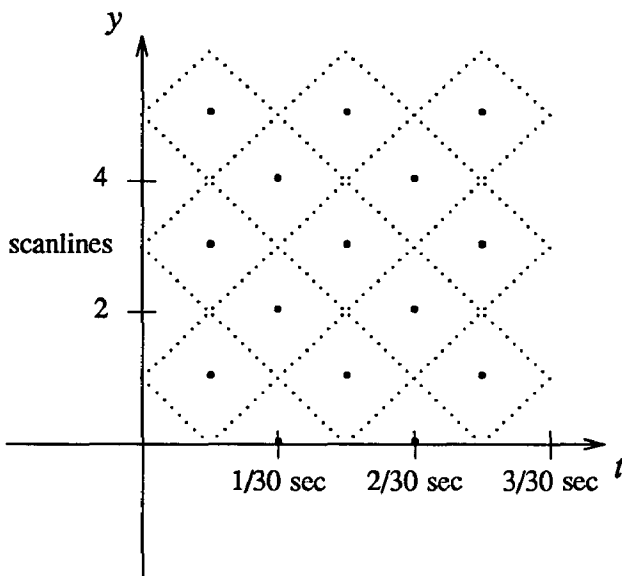Another interesting possibility is a *spatiotemporal diamond* shape, shown in Figure 7.

Figure 7. Spatiotemporal Diamond Filter

Altho this filter has a pleasing symmetry, the resulting video has serious problems. Again, there is too much motion blur, and too much vertical sharpness which causes interline flicker.

In fact, for the rotating-carousel animation, the best results were obtained by doing no motion blur at all. The filter used was like the interlaced box, but with no temporal width. At 10 r.p.m., motion judder was not apparent (even viewing digital video on studio monitors), and the image appeared noticeably sharper when no motion blur was used.

Even without motion blur, it is still vitally important to

generate samples that are interlaced in time. This is well known to animation experts [Chuang85]. The enlarged portion of Figure 6 shows that a still frame from an interlaced animation has a serrated appearance. However, the moving image looks fine. If the animation was generated without interlacing, the opposite effect is seen; still frames look fine, but the moving image has an unpleasant serrated appearance.

Motion judder is not quite as serious a problem for video as it is in motion pictures, because the temporal sampling rate is higher. In an advanced renderer, an optional degree of temporal antialiasing would still be useful. Rapid period motion (e.g., a spinning wheel or a hummingbird's wings) demands temporal antialiasing. We believe the temporal width of the prefilter should be adjustable, like the shuttle angle in a motion picture camera.

5. Spatiotemporal Antialiasing

An annoying aliasing artifact of interlaced video is interline flicker. This is spatiotemporal aliasing caused by high spatial frequencies near 0.5 CPS (see: Figure 2) being replicated as a 30 Hz temporal frequency. Interline flicker can cause large areas of a display to flicker. However, a more typical manifestation of flicker is crawling jaggies or fluctuating Moiré patterns.

Unfortunately, there is a tradeoff between flicker and vertical sharpness. A filter like the spatiotemporal diamond is too sharp and causes a lot of flicker (as well as too much motion blur). To some extent, this is the fault of current display technology, as well as a problem with interlacing itself. Better reconstruction of the video signal might remove flickering effects by postfiltering to remove the offending sideband. But with current interlaced monitors, reduction of vertical sharpness is the only remedy.

Moving jaggies seem to be most noticeable on near-horizontal edges moving vertically at a slow rate (around 30 scanlines per second). However, even when objects are stationary, flicker can be annoying. Figure 8 shows a side-by-side view of the even and odd scan lines of a frame of a stationary zone plate, prefiltered with a box filter only one scanline high. On an interlaced display, these two complimentary Moiré patterns alternate at 30 Hz. The image shimmers when viewed with a steady gaze, and blinking or head movement cause the Moiré pattern to appear vividly.

A more common (and less extreme) example of still images exhibiting interlace problems is *line crawl*. Moving jaggies can appear, even at the edges of a still image. This is often seen in the output of poor-quality

character generators.

Some have suggested that strong vertical filtering be used when there is movement, and greater vertical sharpness allowed for still images [Chuang85]. We are not convinced that even this is safe, given that a still image with too much vertical sharpness can still exhibit line crawl and flicker as Figure 8 shows.

Of the temporal filters described so far, the interlaced box (Figure 4) was the best. This filter is two scanlines high, and its vertical frequency response (i.e., Fourier transform) is a sinc function with a zero at the offending frequency of 0.5 CPS.

Smoother filters exist which have a notch in their spectrum at 0.5 CPS and attenuate other high frequencies better than a box filter. A family of smooth piecewise-cubic filters has been reported by Mitchell and Netravali [Mitchell88], and one member of that family has a notch at the desired frequency. In fact, this particular member of the family happens to be quadratic:

$$f(y) = \begin{cases} -0.25\,|\,y\,|^2 + 0.5 & \text{if } |\,y\,| < 1 \\ 0.25\,|\,y\,|^2 - |\,y\,| + 1.0 & \text{if } 1 \le |\,y\,| < 2 \\ 0 & \text{otherwise} \end{cases}$$

The support of this filter is four units in width. One way of using it would be to sample a region of the image four scanlines high and weight the samples with this filter. A similar result is obtained by unweighted sampling within a *shaped aperture* with a width proportional to $f(y)$ (as in Figure 9).

This is reminiscent of the use of shaped apertures in early facsimile scanning machines [Mertz34]. As in those machines, the vertical frequency response of the aperture is designed to reduce aliasing. However, this aperture is not continuously moved across an image. Instead, each successive pixel value is derived by integrating within a corresponding fixed aperture, as shown in Figure 9. These apertures are two pixels wide (i.e., 4:2:2 standard pixels) at the center. This is done to ensure better horizontal filtering and to suppress chroma aliasing (discussed in the next section).

In Figure 10, the even and odd scan lines of a zone plate are shown using this notch filter. The flickering Moiré patterns are nearly absent.
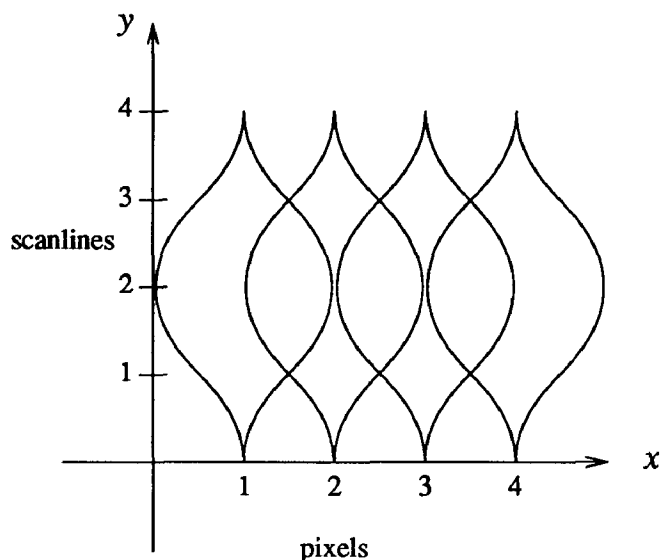

Figure 9. Notch-Filter Apertures

6. Chroma Aliasing

No matter what type of video format a computer-generated animation is created on, it will almost certainly be converted to *composite* video for broadcast or viewing on conventional home television sets. To maintain compatibility with older systems, composite video consists of a *luminance* (black-and-white) video signal with low-bandwidth *chrominance* (color) information encoded into a portion of the spectrum which usually contains little important information [Dubois88].

Figures 2 and 3 show the approximate location of the color subcarrier, which is modulated with the chrominance signal (consisting of two superimposed color signals 90 degrees out of phase). From Figure 3, it is clear that if the luminance signal contains very high diagonal frequencies, it may overlap with the chrominance signal. The resulting chroma aliasing can be seen in Figure 11 as two colored bullseye patterns on the left and right sides (the aliasing at the top center is spatiotemporal).

Chroma aliasing occurs at a later stage than spatial and temporal aliasing. It is created by the electronics that encode and decode composite video signals. If the best encoding hardware is used, it is not generally a problem (even in the test pattern shown in Figure 11), but such hardware is expensive. The encoding that occurs in typical video recording equipment does not prefilter the video carefully enough before forming the composite signal.

The notch aperture used to suppress interline flicker also does a fairly good job of reducing chroma aliasing. Because the filter is somewhat diamond shaped (see

Figure 9), it tends to attenuate diagonal frequencies. Chroma aliasing in the zone plate image was greatly reduced in comparison with the interlaced box filter.

7. Conclusions

Generating synthetic interlaced video presents a number of problems not encountered when synthesizing a motion-picture animation. Interlaced scanning leads to the dilemma of having to choose between vertical sharpness or flicker. The color encoding of NTSC allows chroma aliasing if a scene contains high-frequency diagonal structure. In this paper, we describe some experiments with video synthesis and present an approach used in our multiformat rendering system.

Interlace flicker is a problem in still images as well as scenes with motion. Better "de-interlacing" display systems could help someday, but all that can be done now is to deal with issues in the video source. This means reducing the vertical sharpness by suppressing frequencies around 0.5 cycles per scanline. Video synthesis software is not constrained to imitate television cameras, but our experiments with highly unconventional schemes, such as the spatiotemporal diamond filter, did not alleviate the flicker problem.

Chroma aliasing and flicker are both treated by area averaging of pixels over a shaped aperture. The shape of this aperture is selected so that its frequency response will have a notch at the vertical frequency of 0.5 CPS, and so that high diagonal frequencies will be suppressed.

Temporal aliasing is a problem for both video and motion pictures. The tradeoff between motion blur and motion judder is complex and depends strongly on whether or not viewers track a moving object with their eyes. Temporal filtering should be an option which the artist can adjust. It is clear that aggressive low-pass filtering is not appropriate in the temporal domain, which is why sophisticated filter designs have not been considered. A possible future consideration is to allow the artist to apply different amounts of temporal filtering to different objects in the scene.

8. Acknowledgements

We would like to thank Arun Netravali and Don Pian for many useful discussions about the theory of television. Many thanks to Gin Qua for his help in preparing our video presentation. Thanks to Pat Hanrahan and the SIGGRAPH reviewers for their comments.

9. References

[CCIR601]   "Encoding Parameters Of Digital Television For Studios", CCIR Recommendation 601-1.

[Chuang85]   Chuang, Richard, "Rendering for Television", SIGGRAPH Tutorial, 1985.

[Cook84]   Cook, Robert L., Thomas Porter, and Loren Carpenter, "Distributed Ray Tracing", SIGGRAPH 84, July 1984, pp 137-145.

[Dubois88]   Dubois, Eric, and William F. Schreiber, "Improvements to NTSC by Multidimensional Filtering", SMTPE Journal, June 1988, pp 446-463.

[Hsu85]   Hsu, Steve C., "Motion-Induced Degradations of Temporally Sampled Images", Master's thesis, MIT Department of Electrical Engineering, June 1985.

[Murata83]   Murata, Nobuo, et al., "Development of a 3-MOS Color Camera", SMPTE Journal, December 1983, pp 1270-1273.

[Mertz34]   Mertz, Pierre, and Frank Gray, "A Theory of Scanning and Its Relation to the Characteristics of the Transmitted Signal in Telephotography and Television", Bell System Technical Journal, July 1934, pp 464-515.

[Mitchell88]   Mitchell, Don P., and Arun N. Netravali, "Reconstruction Filters in Computer Graphics", SIGGRAPH 88, August 1988, pp 221-228.

[Potmesil83]   Potmesil, Michael, and Indranil Chakravarty, "Modeling Motion Blur in Computer-Generated Images," SIGGRAPH 83, July 1983, pp 389-399.

[Potmesil89]   Potmesil, Michael, Eric M. Hoffert, "The Pixel Machine: A Parallel Image Computer", SIGGRAPH 89, July 1989, pp 69-78.

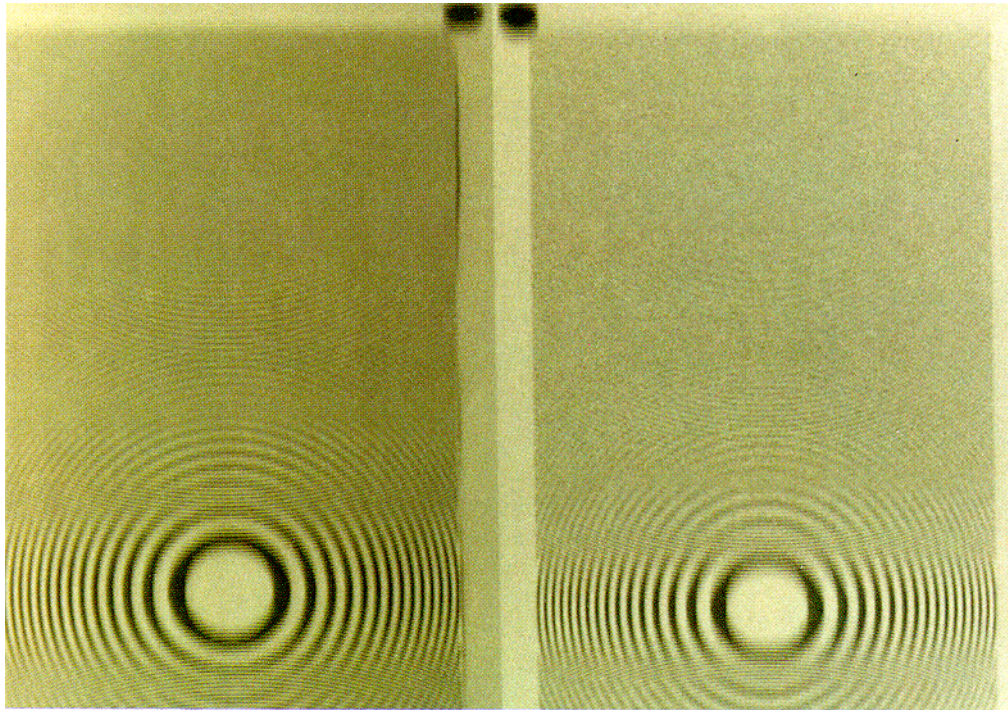[Tonge84]   Tonge, G. J., "The Television Scanning Process", SMPTE Journal, July 1984, pp 657-666.

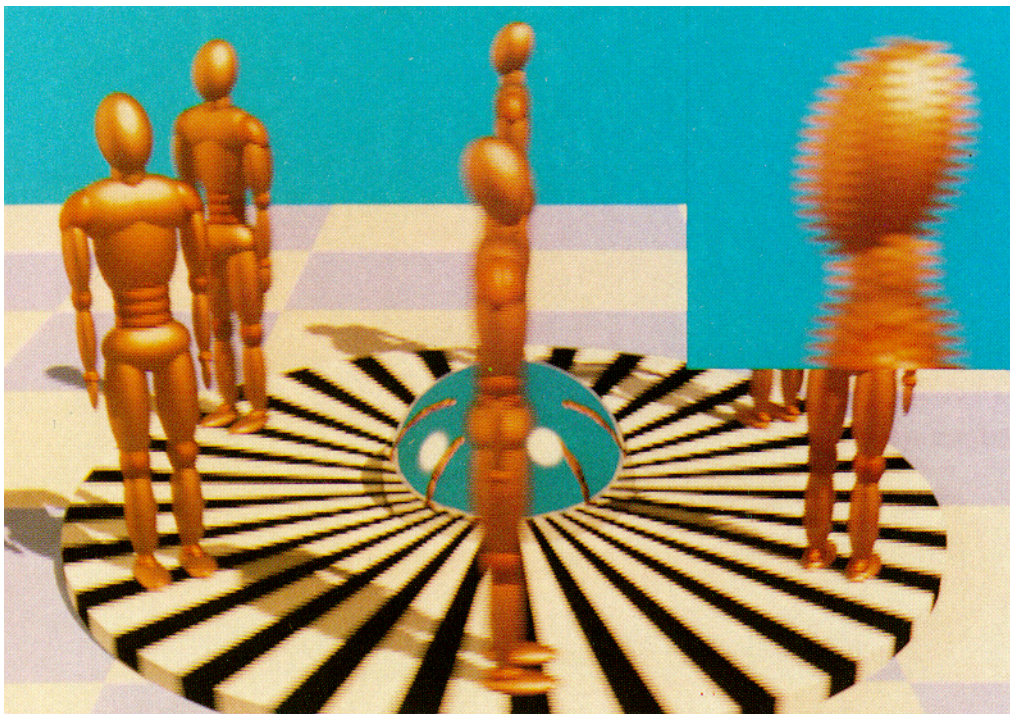Figure 5. CCD Camera and Software Simulation
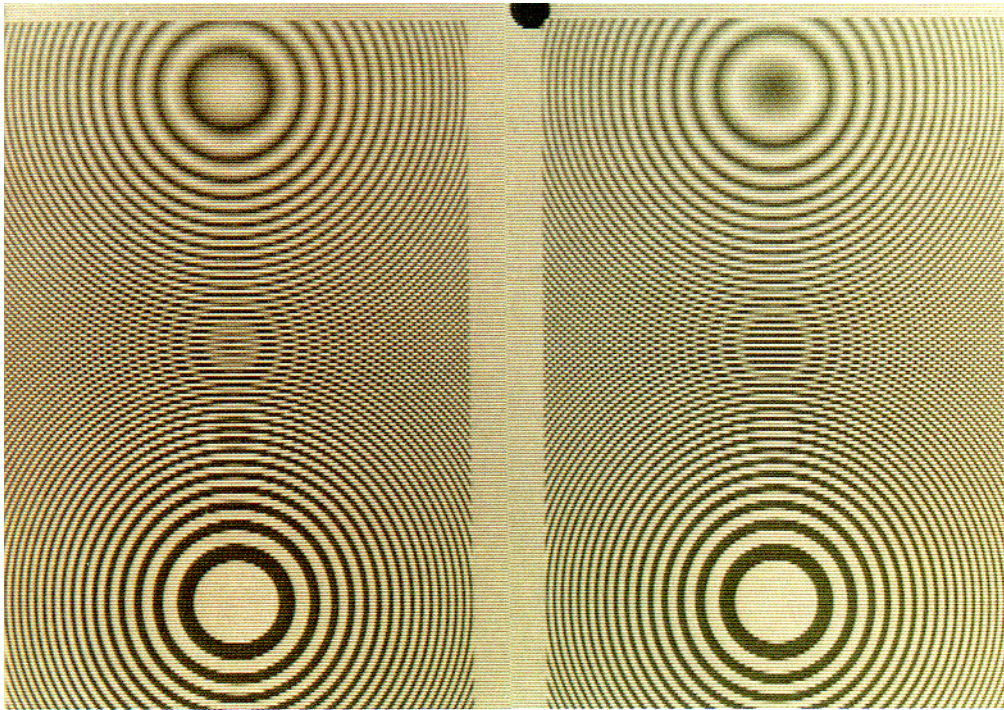
Figure 6. Video Frame from an Animation

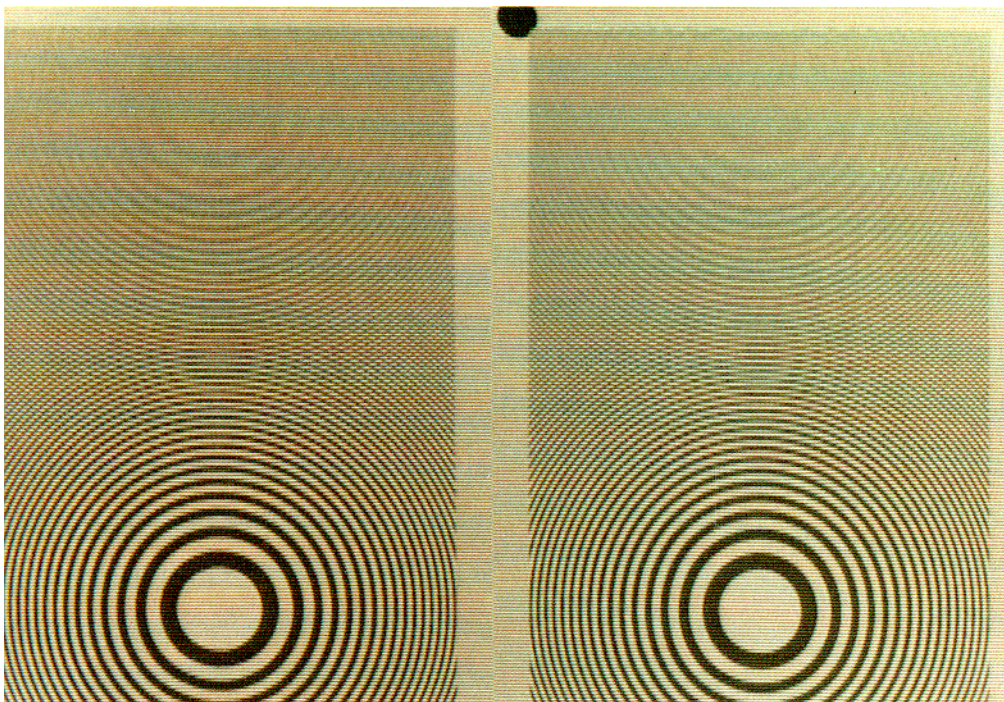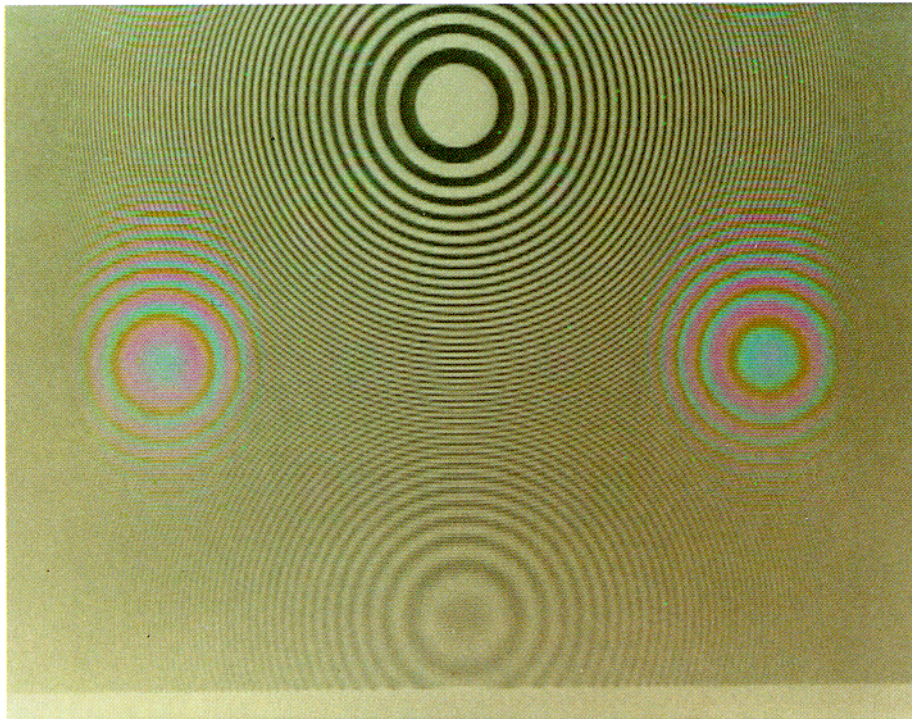Figure 8. Flickering Moiré Pattern (odd & even fields)

Figure 10. Notch-Aperture Filtered (odd & even fields)

Figure 11. Chroma Aliasing