What is the resolution limit for high quality video?

Abstract

How high does video quality need to be when it is viewed on a television, digital tablet or smartphone? For these displays, moving images are converted to a fixed-pixel grid, which can impact the perception of moving images. But at what resolution levels do viewers begin to notice "jaggies" that indicate a noticeable decline in video quality? And beyond what resolution does there cease to be an improvement in video quality?

While most users are familiar with some of the traditional static acuity measurements such as the Snellen eye chart, for video we are concerned with a "dynamic" acuity type of measurement. That is, when moving objects are displayed, at what pixel density do we no longer notice any artifacts caused by the discrete nature of the pixels? This paper presents the results and findings of research conducted by Pixelworks to identify pixel density perception limits and the corresponding optimal level for display resolution.

Background

Beginning with the introduction of the Apple iPhone 4 in June 2010, displays whose individual pixels can't be perceived have been considered the gold standard for phones and tablets. For the iPhone 4, this display was 57 pixels per degree (ppd), or 326 pixels per inch (ppi) when viewed from a 10" distance. While some people disagreed(1), in general this measurement was accepted as correct.

The basis of this acceptance was research initially done by Dr. Hermann Snellen—creator of the familiar Snellen Chart for vision tests—that indicated the threshold for recognizing letters was 5 arc minutes. The letters on the eye chart can be represented with as few as 5 pixels in the vertical direction, which is the equivalent of 60 ppd for individuals with 20/20 vision.

However, vision tests are not performed with displays that actually contain pixels. While not important for static images, when objects move, as they do on video, the pixels that display them will remain in the same place, which means the content on the display will change. For video, an unanswered question is whether this change will be visible to the naked eye.

Specifically, what are the equivalent pixel-perception limits for moving images that are both captured and displayed by a fixed-pixel device? After all, the content of an eye chart looks nothing like the typical scene in a movie displayed on a fixed-pixel device. Through a discussion of simulations and experiments conducted by Pixelworks, this paper will try to answer these questions to determine the pixel density limit for moving images.

	Viewing Distance	Screen Size	Horizontal Resolution	Vertical Resolution	Pixels Per Inch	Pixels Per Degree
iPhone 5	10''	4''	1136	480	308	54
Galaxy 4	10''	5''	1920	1080	441	77
HTC one	10''	4.7''	1920	1080	469	82
iPad	15''	10.1''	1024	768	127	33
iPad Retina	15''	10.1''	2048	1536	253	66
Next Gen Kindle	15''	9''	2560	1600	335	88
MacBook Pro	20''	13.3''	2560	1600	227	79
Sony Fit 14	20''	14''	1600	900	131	46
MacBook Pro	20''	15.4''	2880	1800	221	77
FHD @ THX dist	67''	56''	1920	1080	39	46
FHD @ SMPTE dist	90''	56''	1920	1080	39	62
FHD @ Man dist	140''	56''	1920	1080	39	96
4K @ THX dist	67''	56''	3840	2160	79	92
4K @ SMPTE dist	90''	56''	3840	2160	79	124

Table 1: Parameters of common consumer devices for viewing video

Display Quality and Video Artifacts

Table 1 lists a variety of devices used to view video, from smartphones to TVs. Most cutting-edge devices with displays have resolutions that are equal to or greater than 60 ppd, more than the iPhone 4. With this level of display quality available today, the question becomes whether additional increases in display resolution are worth the additional expense.

What types of artifacts will viewers see when using these devices to view video? After all, video does not consist of only static letters or frequency gratings; it also contains textures of multiple frequencies and different background levels.

To understand the impact of textures, think of a coarse cloth worn by somebody who is not perfectly still. Would you notice that the cloth seems to vary in brightness or sharpness as it moves, even though the motion is too small for changes to be caused by motion blur? While these changes may not be objectionable, are they enough to be perceived? That is, are these changes enough to provide visual clues that you are not watching the performance in person?

The concern is not whether you can recognize the content or orientation of an image, as in the case of the Snellen Chart. Instead, the concern is whether you can notice that something looks different if it moves by a fraction of pixel, differences that would be perceived as unnatural.

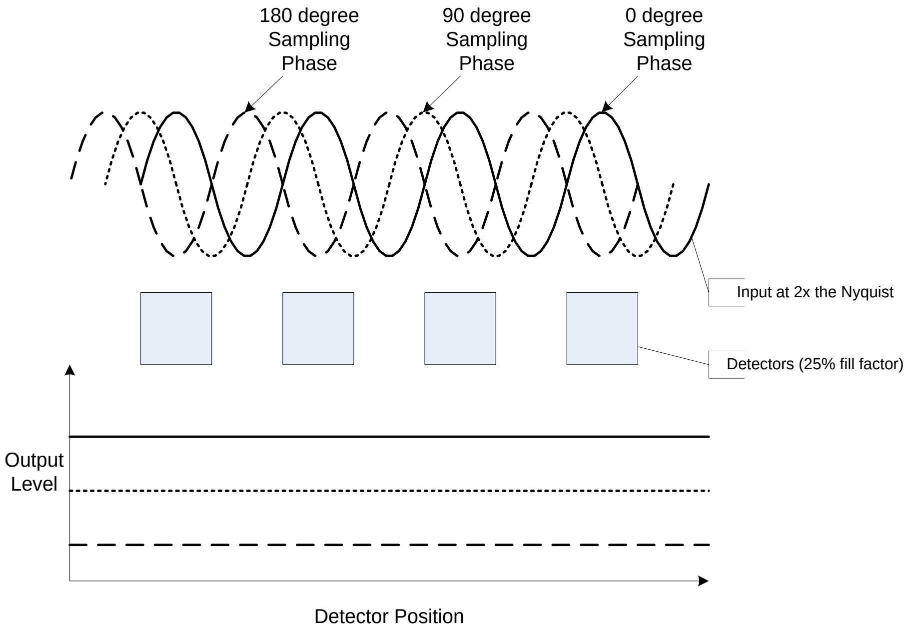


Figure 1: Changes in pixel value caused by sampling phase

Sampling Phase, Aliasing and the Ability to Notice Jaggies

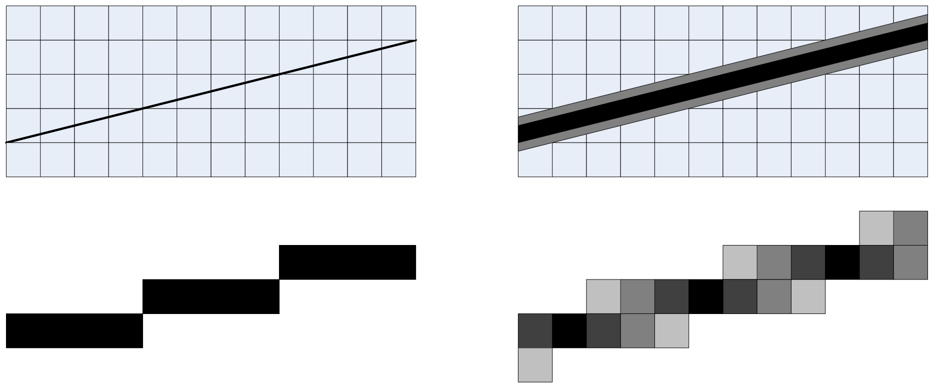
One limitation of previous studies is that they are based on targets generated on devices with significantly higher resolution than the best consumer displays. Furthermore, when viewed on TVs, tablets or phones, video has moving images that have been converted to a fixed-pixel grid. This display difference implies that we should consider both resolution and the sampling phase of the fixed-pixel grid relative to the viewed object when determining whether a viewer can notice the fixed-pixel nature of the display.

Also, because of the resolution difference, it may not be enough to simply apply the resolution limits typically quoted in the literature to video images. The issues of sampling phase, aliasing and the viewer's ability to notice "jaggies" are also important to consider.

The first issue is the impact of sampling an object that is not constrained to integer pixel positions. Specifically, is it possible to detect that an object is covering a varying number of pixels as it moves across the display?

To illustrate this point, let's use an extreme case where the light-sensitive portion of a pixel is very small compared to the spacing between pixels (a low fill factor) and the input frequency is high enough to cause aliasing. The aliasing causes the higher frequencies to fold back into the output at lower frequencies, which are more visible.

For example, if the input were a sine wave at the sampling frequency, then every pixel would have the same value, but that value would depend on the sampling phase. That is, motion will cause the amplitude of a low-frequency signal to change and that change



a) Small Optical Spot Size

b) Spot Size Approx. Equal to Detector

Figure 2: Aliasing perceived by a viewer when the optical system and detector are not matched

would be noticeable to a viewer regardless of the display resolution. (Figure 1)

While Figure 1 presents an artificial pattern, even natural content can be impacted if the optical system and detector are not matched. For example, a black-and-white edge or a line that is not perfectly vertical or horizontal could have noticeable jaggies if the spot size is too small. If the spot size is significantly smaller than the pixel spacing, the diagonal edge will shift from one pixel to the next.

The ability to see jaggies is related to an existing measurement called vernier acuity. The literature indicates that the vernier acuity is 10 arc sec (10”) or 1/6th of the static acuity limit for a pixel. Therefore, if the spot size is too small, the diagonal line will be perceived as a series of small line segments (jaggy edges) even if you cannot actually see the pixels. (Figure 2)

To simulate the impact of the display resolution with respect to vernier acuity, a model was constructed to simulate an edge. To simplify the calculations, the model assumes a square spot size with a box-car distribution equal to the detector size. The model then uses a line-spread function (LSF)-based, low-pass filter corresponding to a pupil diameter of 2.4mm[2], which is close to the pupil size found for television viewing(2). The difference in position between the two objects on the retina is measured at the 50% point.

Figure 3 shows the worst case for a detector misalignment to the edge of approximately 0.25 pixels. When the differences in the edge position are measured after the eye’s optical low-pass filter is applied, the difference is less than 5” with a 60 ppd display. This result suggests that as long as the optical system is correctly designed at the capture side, the jagged edge of diagonal features should not be noticeable.

The discussion in the rest of this paper assumes the capture device has limited the optical resolution at the sensor to less than one-half the device’s pixel density, meaning there is little to no aliasing in the capture process.

Determining Resolution with Optical Gratings

Optical gratings are often used to measure the resolution of the eye in ophthalmological studies. However, the type of gratings used in these studies is a special case, where motion is not expected to impact the results. In this case, no difference is expected for fixed-pixel devices because the ability to

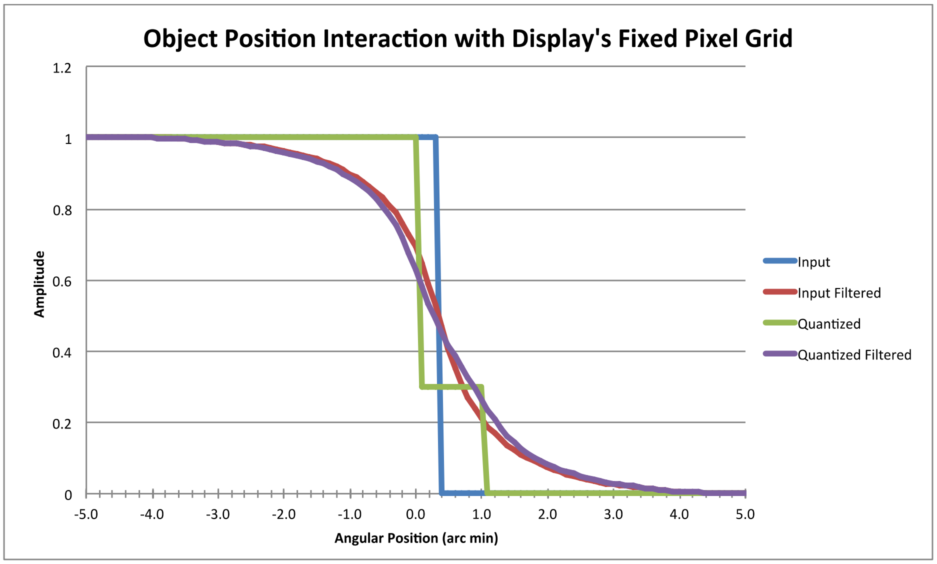


Figure 3: Results for worst-case detector misalignment

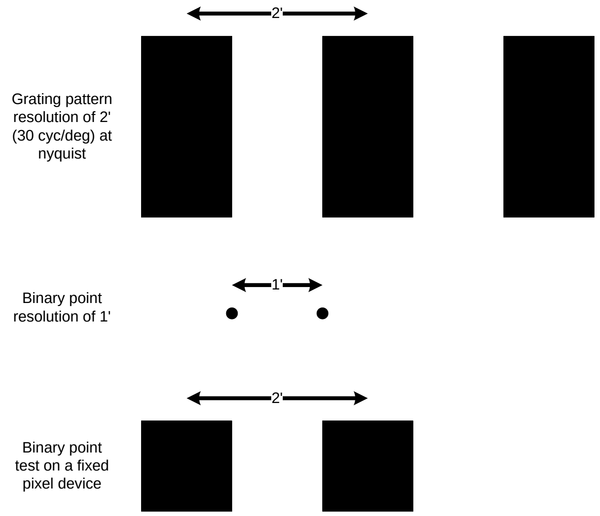


Figure 4: Grating pattern density on a fixed-pixel device

notice the pixels for that type of pattern is already being compared to what the pattern looks like when sampled ½ pixel out of phase: a flat gray field.

But what about the resolution results when viewing other content? The literature suggests that the ability to resolve two points is twice that of a grating, e.g., if the distance between two points is equal to 1', the distance between the centers of two white bars is 2'.

However, the point source size must be the same size as a pixel. For a fixed-pixel device, it is not feasible to generate a point source significantly smaller than the distance between the pixels(3) and still have a reasonable brightness level. To estimate the limit for a binary point-source target we need to treat the actual pixel size as if it was a collection of smaller points of light. When we do that, we find that the 1' separation is actually the distance between the edges of the pixels. For a fixed-pixel device, the effective pixel density is the same for both binary point and grating targets at 30 cyc/deg or 60 pixels/deg (ppd). (Figure 4)

3D Depth Perception

The final static performance measurement related to video content is 3D depth perception. The ability of the eye to notice depth differences under ideal conditions translates to acuity of 5", twelve times that of resolving an individual pixel. As shown in the earlier discussion about vernier acuity, 60 ppd is good enough for 3D depth perception when there is adequate optical low-pass filtering at the capture side.

Test Pattern for Testing Moving Images

Real image content consists of more than stationary single edges, point sources or grating patterns. Consider two closely spaced lines such as branches on a tree. At what point can a viewer detect that they are two separate lines? If the lines are moving, does the sampling phase change the ability to detect them as two separate lines? If this were the case, motion would cause an object to show a variation in sharpness depending on the sampling phase. Although you may not be able to see the pixels, you would see the image quality of the object change as it moves and perceive that change as unnatural.

Unfortunately, as will be discussed in a later paper in this series, it is very difficult to construct and display content that will be seen exactly as if the content of one pixel has been spread over two pixels (sampling phase = 180 degrees) and that is reproducible on multiple displays. To address this difficulty, instead of creating a pattern at the native pixel density of the display, patterns are created at twice the size.

Having patterns at twice the size allows replacing the lower amplitude, two pixel-wide, pattern when the sampling is at 180 degrees out-of-phase with a halftone pattern. The halftone pattern will be blurred enough by the eye to be indistinguishable to a gray-tone image because the subject will be at 2x the distance of where you can distinguish a pixel because the test is now about distinguishing patterns that are at least 2 pixels in each direction. Using two-pixel wide objects also allows a change in sampling phase while keeping the objects spatially aligned so that results will not be confused with detecting a change using the person's vernier acuity or be confounded by the display technology (motion blur).

In a test of moving images, three different patterns were used: grating, two-line stair step and single point (pixel). The halftone and solid portions of the pattern were switched at a 1 Hz rate. (Figure 5) The task for the subjects was to be able to identify which portion of the test pattern had the

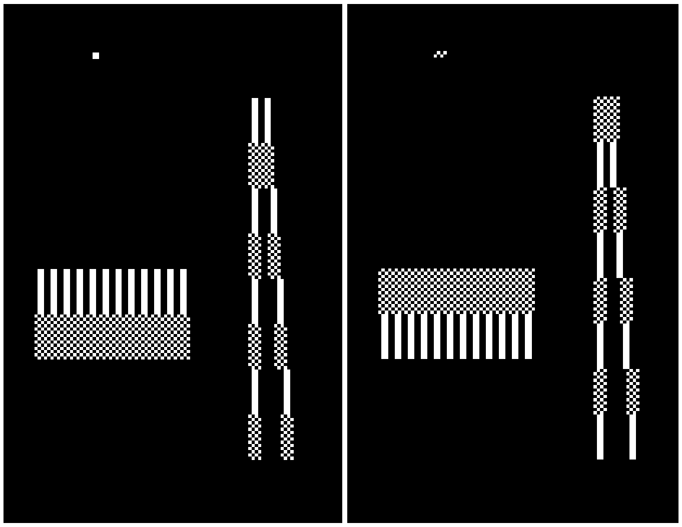


Figure 5: Original and switched testing patterns

halftone (less sharp or different thickness) vs. the solid black and white sections. Unfortunately, the simpler and more realistic task of just asking the test subjects to notice that the display has changed is confounded by the response time of an LCD display.

The two-line test presented a stair-step pattern of one-pixel-wide lines (two pixels on the display) separated by different distances (1x, 2x, 3x and 4x the line width) along with a section representing the sampling phase shifted by ½ pixel. (Figure 5) Preliminary testing showed that the subjects were most sensitive to a two-line pattern with the 2x line-width separation between the lines. Because spatial frequencies can also play a role in the ability to distinguish objects, that portion of the stair-step pattern was repeated with different vertical frequencies between the sampling phases. (Figure 6)

Testing also showed that the overall size of the pattern played a role in the subjects' perception. Two patterns were added that extended the tests to multiple lines so the areas subtended by all patterns were approximately the same. (Figure 7)

The patterns were embedded on a gray background to provide an average illumination level that is similar to watching video content. The test patterns were repeated for white-on-black and black-on-white perception. The test was administered to individuals within Pixelworks who work on image quality and should be considered trained observers.

Detecting Results for the Moving Image Pattern Testing

Overall, higher than expected sensitivity to the patterns was observed in the test subjects. For example, the average detection distance for the grating pattern corresponded to an angular resolution of 78 pixels/degree (Table 2) versus 60 ppd in the literature. This result could be due to changing the test condition from adjusting contrast(4) until the grating is visible to trying to detect a difference between flat gray and a switching pattern. Further testing is required to see if the increase in sensitivity is the result of testing that more closely represents motion or a bias in the acuity of the test subjects.

As shown in Table 2, the most obvious impact on acuity is the size of the testing pattern. The lowest acuity corresponds to the single-pixel object and the highest acuity is for the multi-line vertical frequency pattern. In fact, for objects that are close in size to the

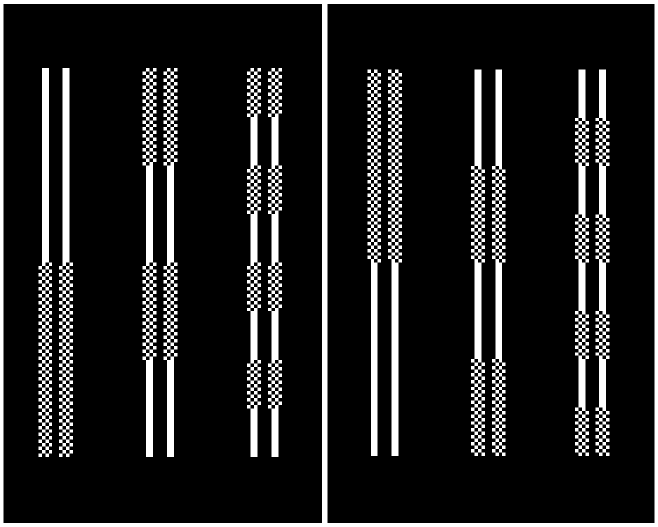


Figure 6: Stair-step patterns for moving image testing

standard Snellen chart letters (e.g., the two-line step), the acuity is close to 20/20 (61 pixels/degree).

The high standard deviation value for the multi-line step pattern vs. the grating pattern was unexpected given that the size of the pattern and the method of detecting acuity was the same. Analysis of the raw data shows that even though the multi-line step pattern included the grating pattern, three individuals perceived a lower resolution threshold for the multi-line step pattern vs. the grating pattern. Because of this difference, the results from those three individuals were eliminated from all measurements and the statistics recomputed as shown in Table 3.

After eliminating the three outliers, the difference in acuity between the grating pattern and the multi-line step pattern becomes significant, with a confidence level of 99.98%(5).

Discussion of the Detection Results

In order to understand the results, the two-line pattern with 2x and 3x separation between the lines was simulated. The change in modulation depth was then compared between the case when the pattern is aligned on the pixels and when it is shifted by ½ pixel.

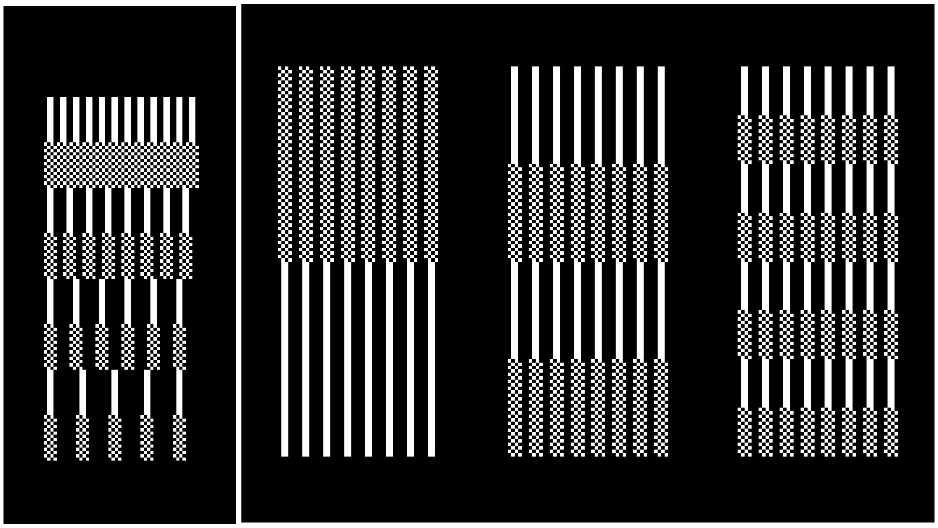


Figure 7: Extended testing with multi-line patterns

In the calculations, the modulation depth is the peak-to-peak divided by 2x the average, which provides the same result as the literature when used for a grating pattern. The LSF used to model the eye's optical response was obtained from a previously published study so that the tails of the low-pass filter (LPF) kernel would be modeled accurately. The LSF is based on a pupil diameter of 2.4mm(6), which is consistent with viewing TV in a home environment.

Figure 8 shows a computer simulation of the two-line pattern separated by 2x the pixel width on the screen at the approximate angular resolution (100 ppd) where most subjects find the two sampling phases barely distinguishable. The 0 pixel offset line pair has a modulation depth of 33% and the ½ pixel offset line pair has a modulation depth of 15%. The ½ pixel offset has reduced the modulation depth by 55%.

If we now look at the section of the stair step where the distance between lines is 3x the line width, with the test subject at the same distance, the modulation depth changes to 42% for the 0 pixel offset and 28% for the ½ pixel offset. While the modulation depth is larger, the reduction caused by the ½ pixel offset is now only 33%, supporting the finding that motion acuity is maximized with a gap between lines of exactly two pixels. (Figure 9)

Pattern	Average (pixels/deg)	Standard Deviation
Single Pixel	47	9.2
Grating	78	7.5
Two Step	61	9.6
Two Line Vert. Freq.	81	14.0
Multi-line Step	82	13.8
Multi-line Vert. Freq.	88	8.3

Table 2: Detection results from pattern testing

Pattern	Average (pixels/deg)	Standard Deviation
Single Pixel	48	9.2
Grating	78	7.9
Two Step	61	8.7
Two Line Vert. Freq.	81	15.1
Multi-line Step	88	8.7
Multi-line Vert. Freq.	90	6.1

Table 3: Recomputed detection results

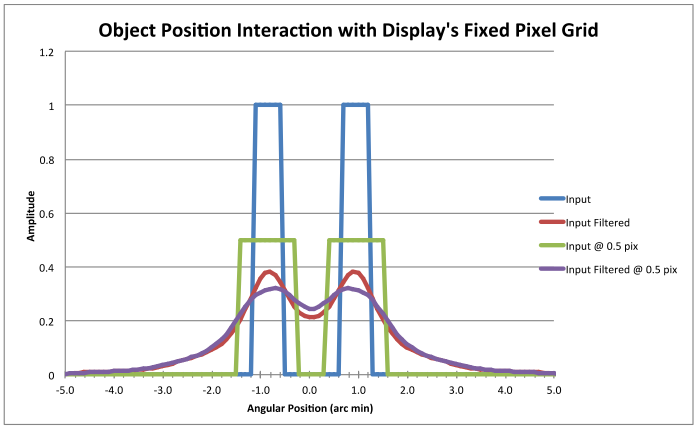


Figure 8: Modulation depth for a two-line pattern with 2x pixel-width separation

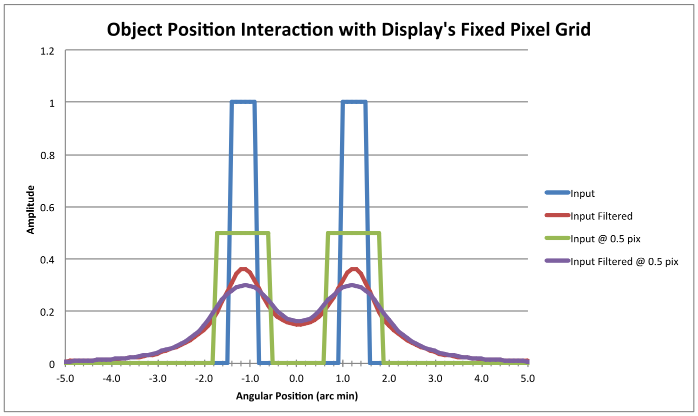


Figure 9: Modulation depth for a two-line pattern with 3x pixel-width separation

Conclusion

These experiments show that when motion is taken into account, standard resolution tests do not directly relate to the limits of perception. Specifically, when determining the resolution limit for moving images on fixed-pixel devices, the input sampling phase and the size of the pattern need to be taken into consideration.

When the sampling-phase effects are taken into account, the perceivable pixel density limit increases from 60 ppd to nearly 90 ppd. Furthermore, differences from perfect reproduction of the captured scene on the display device can lead to the situation where you will always see the impact of the fixed-pixel capture and display.

Common video processing techniques can cause the fixed-pixel structure of the display to be visible regardless of the display resolution. Fortunately, alternative approaches are available that can greatly reduce the effect of fixed-pixel sampling, as will be discussed in the next paper in this series.

Footnotes & References

(1) “Analyst Challenges Apple's iPhone 4 'Retina Display' Claims”, http://www.pcmag.com/article2/0,2817,2364871,00.asp

(2) The pupil size for television viewing in a dark room can be calculated using the following steps:
  a. Use the illumination level of the display or screen and the average 40% picture level of video.
  b. For televisions, the maximum illumination is typically 500cd/m2, which implies an average illumination of 200cd/m2. From figure 36.21 in the Handbook of Optical Systems (Herbert Gross, Handbook of Optical Systems: Vol. 4 Survey of Optical Instruments, Chapter 36, 2008) we find a pupil diameter of 2.5mm.

(3) Of course, we could limit the test to just the green channel, with the green subpixels separated. However, the goal is not to construct a static binary point test but to see if there are issues with subpixel motion on video displays.

(4) It is possible that if the contrast is adjusted slowly enough, the sensitivity to the pattern is reduced by the inhibition of the eye to the current pattern.

(5) 1 sided T test

(6) F. W. Campbell and R. W. Gubisch, “Optical Quality of The Human Eye”, J. Physiol, p. 571, 1966, v186