



Ver 2.0 - Feb 9, 2006



## What is a Video Processor and Why is it Important?

High-definition video is a wonderful thing. Who couldn't love the beautiful picture and digital surround sound? Who wouldn't trade their old 27" or 32" CRT for a beautiful new HDTV? It's time to say goodbye to low resolution and image speckles from poor-quality NTSC signals once and for all. Get that pristine HD image and life will be great, right?

Not so fast. Sure, you can now view HD signals, and they look really great. But what happens when you feel like watching one of your favorite old TV shows that isn't in HD? All that noise and garbage in the image is not gone; instead, it's accentuated! Many people are shocked to see that the picture on their new HDTV looks worse than their old TV when watching their favorite shows!

Unfortunately, few people realize when they purchase their first high-definition TV set that the majority of TV programming is still delivered as standard-definition (SD) 480-line interlaced video through composite, S-video, or component connections from cable TV set-top boxes, VHS VCRs, DVD players, and Mini DV camcorders. That SD image must be enlarged to fill the big screen on that new HDTV, and all the flaws in interlaced SD images are enlarged along with everything else.

In contrast to conventional CRT televisions, fixed-pixel displays dominate today's home theater marketplace, from single-chip and three-panel front projectors to large, flat-panel direct-view monitors. There are plenty of acronyms to identify the technologies behind these competing products, including LCD (liquid-crystal display), DLP (Digital Light Processing™), LCoS (liquid crystal on silicon), and PDP (plasma display panel).

Each of these technologies creates electronic images in a different way, but they all share one characteristic: a fixed matrix of imaging pixels. This fixed-pixel structure determines the physical resolution of the display and is typically referred to as the display's "native resolution."

In order to convert all incoming video signals to the native resolution of a particular fixed-pixel display, manufacturers must incorporate a video-processing chip inside the display. In addition to scaling the image to fit the native resolution, this video processor is normally designed to enhance the image and remove artifacts caused by the conversion and transmission of video. Surprisingly, the video processor can make a significant difference in overall picture quality, even with true high-definition program content.

Unfortunately, video-processing technology has not kept up with the picture quality of today's larger and larger HD displays, which magnify the image defects that are caused by poor video processing. And although every high-definition display has a video processor, only the highest-quality processor can retain all of the nuance, detail, and intent of the original source. Since video processors can range from a rudimentary \$10 chip used in the lowest-end products to \$70,000 refrigerator-sized boxes used in Hollywood production houses, the video processor plays the most significant role in picture quality today.

## De-Interlacing

Most video sources, including DVD, standard-definition TV, and 1080i high-definition TV, transmit *interlaced* images. Instead of transmitting each video frame in its entirety (what is called progressive scan), most video sources transmit only half of the image in each frame at any given time. This concept also applies to recording video images: video cameras and film-transfer devices record only half of the image in each frame at a time.

The words "interlaced" and "progressive" arise from the days of CRT or "picture-tube" televisions, which form the image of each frame on the screen by scanning an electron beam horizontally across the picture tube, starting at the top and working its way down to the bottom. Each horizontal line "drawn" by the beam includes the part of the picture that falls within the space occupied by that line. If the scanning is interlaced, the electron beam starts by drawing every other line (all the odd-numbered lines) for each frame; this set of lines is called the odd field. Then it resets back to the top of the screen and fills in the missing information, drawing all the even-numbered lines, which are collectively called the even field. Together, the odd and even fields form one complete frame of the video image.

Because all CRT televisions worked in this manner until very recently, the signal transmitted to them was designed to send the odd lines followed by the even lines. This matched the capabilities of the display device, and it also cut in half the amount of information that had to be sent in a given amount of time; in other words, it reduced the transmission bandwidth by a factor of two, which was good news for broadcasters. (Some

modern CRT televisions are capable of scanning each complete frame from top to bottom in a single pass, and thus are said to perform progressive scanning.)

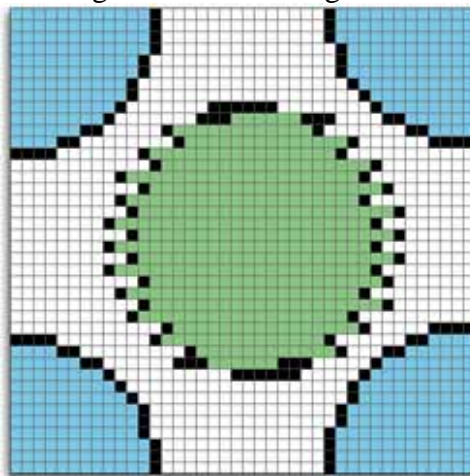
Today, high-definition video displays use digital technologies, such as DLP, LCD, LCOS (including variants SXRD and D-ILA), and plasma, which dominate the television landscape. Instead of "drawing" lines of picture information on the screen, these technologies form images with an array of pixels, and each frame is displayed in its entirety all at once; in other words, all pixels are activated simultaneously to form the complete image rather than forming the image line by line as CRTs do with scanning.

Even so, the video signal that determines what these devices will display is still interlaced or progressive; that is, the information is sent from the source either half a frame or one complete frame at a time. In fact, digital displays ultimately require a progressive signal to operate properly, so if they receive an interlaced signal, it must be converted to progressive before it can be displayed.

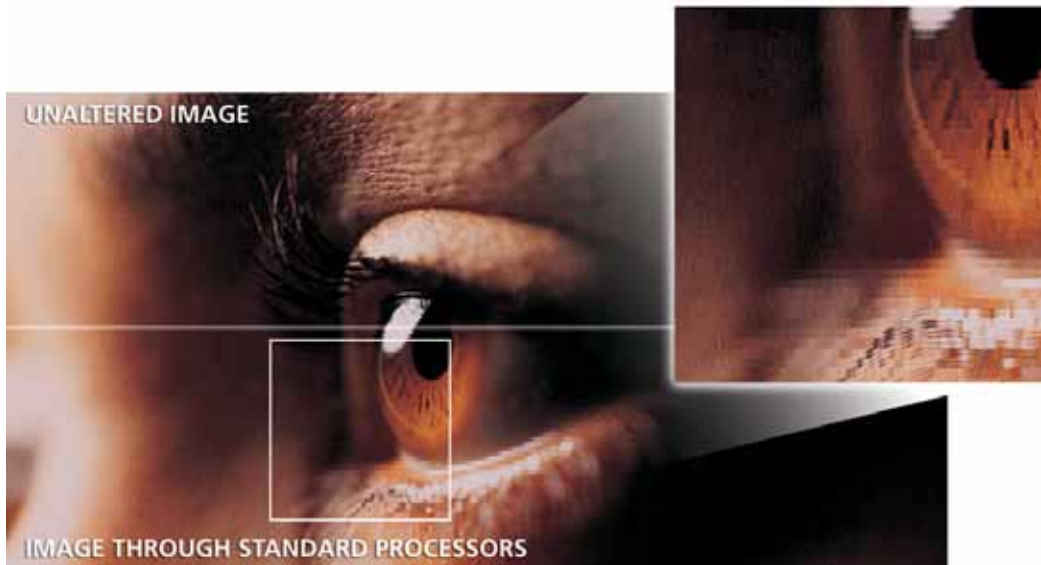
Thus, translating the interlaced video signal from DVD and 1080i sources into progressive format is required by all digital displays. This is the job of a video processor, and the process itself is called *de-interlacing*. Video processors are found in all digital displays as well as many DVD players and other source devices.

If the objects in the video image are not moving, it is very easy to do the de-interlacing – the two fields can be weaved together and combined to form a complete frame. However, if the recording is performed in an interlaced manner, the two source fields that make up a complete frame are not recorded at the same time. Each frame is recorded as an odd field from one point in time, and then as an even field recorded  $1/50^{\text{th}}$  or  $1/60^{\text{th}}$  of a second later.

So, if an object in the video has moved in that fraction of a second, simply combining fields causes the errors in the image called “combing” or “feathering” artifacts.



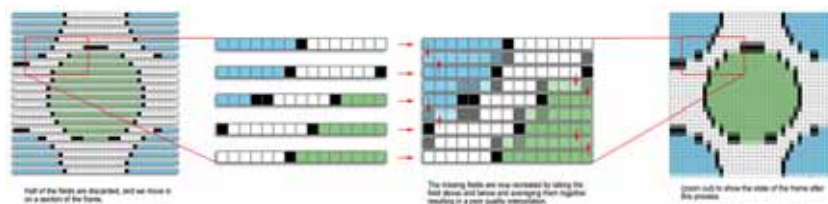
An example of feathering or combing



**Simplest Competitor Approach (Non-Motion Adaptive):**

The simplest approach to avoid these artifacts is to ignore the even fields. This is called a *non-motion adaptive approach*. In this method, when the two fields reach the processor, data from the even fields are completely ignored.

The video-processing circuitry recreates or “interpolates” the missing lines by averaging pixels from above and below. While there are no combing artifacts, image quality is compromised because half of the detail and resolution have been discarded.



More-advanced techniques have been adopted by virtually all standard-definition video processors, but this basic approach is still sometimes used for high-definition signals, due to the increased computational and data-rate requirements of higher video resolution.

With video processors from some competitors, only 540 lines from a 1080i source are used to create the image that makes it to the screen. This is true even for video processors from companies that may have been considered providers of flagship performance in the standard-definition era.

### **Advanced Competitor Approach (Frame-based Motion Adaptive):**

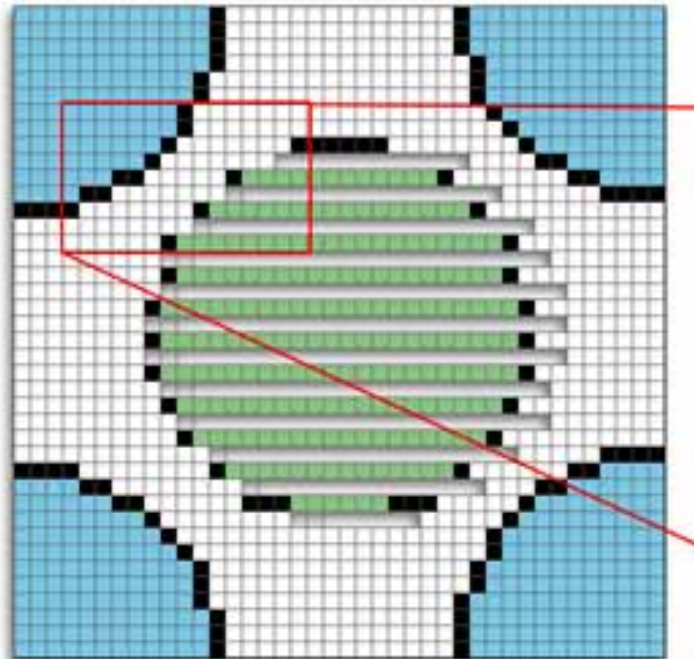
More advanced de-interlacing techniques available from the competition include a frame-based, motion-adaptive algorithm. By default, these video processors use the same technique described above. However, by using a simple motion calculation, the video processor can determine when no movement has occurred in the entire picture.

If nothing in the image is moving, the processor combines the two fields directly. With this method, still images can have the complete 1080 lines of vertical resolution, but as soon as there is any motion, half of the data is discarded and the resolution drops to 540 lines. So, while static test patterns look sharp, video does not.

Frame-based motion-adaptive techniques are now common in standard-definition video processors. However, this is still rare in high-definition video processors due to the computational complexity of even frame-level high-definition motion detection.

### **Silicon Optix HQV Approach (Pixel-Based Motion Adaptive):**

HQV processing represents the most advanced de-interlacing technique available: a true pixel-based motion-adaptive approach. With HQV processing, motion is identified at the pixel level rather than the frame level. While it is mathematically impossible to avoid discarding pixels in motion during de-interlacing, HQV processing is careful to discard **only** the *pixels* that would cause combing artifacts. Everything else is displayed with full resolution.

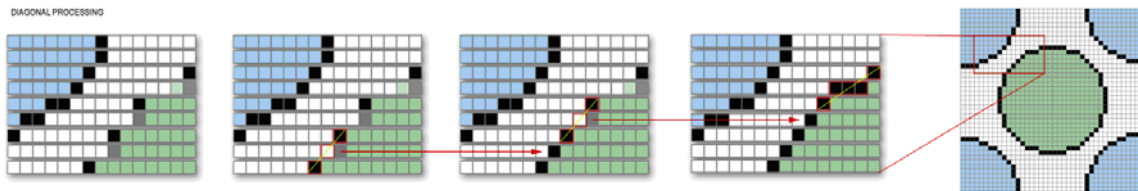


Only the pixels that would cause combing are removed.

Pixel-based motion-adaptive de-interlacing avoids artifacts in moving objects and preserves full resolution of non-moving portions of the screen even if neighboring pixels are in motion.

### “Second Stage” Diagonal Interpolation

To recover some of the detail lost in the areas in motion, HQV processing implements a multi-direction diagonal filter that reconstructs some of the lost data at the edges of moving objects, filtering out any “jaggies.” This operation is called “second-stage” diagonal interpolation because it’s performed after the deinterlacing, which is the first stage of processing. Since diagonal interpolation is independent of the de-interlacing process, competitors have used similar algorithms with their frame-based de-interlacing approaches.



### Truth in Marketing

Silicon Optix is not the only company to implement pixel-based motion-adaptive de-interlacing, and it is important to recognize that all such de-interlacing is not identical. In order to implement a true per-pixel motion-adaptive deinterlacer, the video processor must perform a *four-field* analysis. In addition to the two fields being analyzed in the current frame, the two previous fields are required in order to determine which pixels are in motion. Clearly, if a competing de-interlacer does not evaluate four fields, it simply does not have the data necessary to perform true per-pixel motion-adaptive analysis. Some competing products implement region-based analysis in which motion is determined by evaluating larger blocks of the image rather than complete frames or individual pixels. Obviously, then, a claim of “four-field” analysis alone does not imply per-pixel motion-adaptive de-interlacing.

HQV Processing continues to analyze at the per-pixel level using four-field analysis even in high-definition.

## Film Cadence and Video/Film Detection

Motion picture films are recorded at 24 frames per second. When the movie is released for the home on DVD or a television broadcast, those 24 frames must be converted into 60 interlaced fields. Consider four frames of film: A, B, C, and D.

The first step is to convert these four frames into eight fields. This transforms 24 frames per second (fps) into 48 interlaced fields per second. Then, to account for the faster rate

of the NTSC standard (roughly 30 frames per second or 60 interlaced fields per second), it is necessary to repeat certain fields. This is done by adding an extra field every other frame. That is, both fields of frame A are recorded (A-odd, A-even), but three fields of frame B are recorded (B-odd, B-even, B-odd). The cycle repeats with frames C and D. This is called a 2:3 cadence because two fields of one frame are shown followed by three fields of the next frame.

When this sequence is played back on a progressive-scan video display, it is possible to implement the same de-interlacing techniques described earlier (non-motion adaptive vs. motion adaptive, etc.). However, it is possible to perfectly reconstruct the original frames without losing *any* data. Unlike interlaced video, in which the two fields were recorded a fraction of a second apart, these fields were recorded at the same time in the same film frame and later separated into fields.

So, to display a video signal that originated as 24fps film, all a video processor needs to do is analyze the fields and determine that there is a regularly alternating pattern of two fields followed by three fields, etc. This recognition and reconstruction is called 3:2 pulldown, and it is found in all but the worst de-interlacers. Unfortunately, nothing is quite that simple.

### **Mixed Video and Film**

Sometimes, further editing and post-processing is done on film that has been converted to video. This includes titles, transitions, and other effects. As a result, simply reconstructing full frames results in combing artifacts because parts of the image are best processed using a standard de-interlacing approach, while other parts will look better by detecting the right cadence and reconstructing the original frames.

Like the various approaches to standard de-interlacing, there are many approaches to dealing with mixed video and film. If the processor interprets the material as film, feathering artifacts will appear around the video portion; if the processor interprets the material as video, the film portion will be displayed at half of its resolution. Some processors determine whether there is more film or more video content and choose the approach with the greatest benefit. Since this usually means film, the result is feathering artifacts. Other processors are designed with the idea that these artifacts should never be seen and use the video de-interlacing techniques in all cases, at the expense of as much as half the video resolution.

HQV Processing, on the other hand, uses per-pixel calculations for all of its processing. This means it is possible for the HQV processor to implement cadence-detection strategies for the pixels that represent film content while implementing pixel-based motion-adaptive de-interlacing for the video content that has been superimposed.

# Noise Reduction

Random noise is an inherent problem with all recorded images; the result is often called picture grain. Not only does noise get introduced during post-production editing or the final stage of video compression, but it is also present at the source in the form of film grain or imaging-sensor noise. Noise-reduction algorithms can minimize the grain in a picture.



Image with random noise

The simplest approach to noise reduction is to use a **spatial filter** that removes high-frequency data. In this approach, only a single frame is evaluated at any given time, and parts of the image that are one or two pixels in size are nearly eliminated. This does remove the noise, but it also degrades the image quality because there is no way to differentiate between noise and detail. This approach can also cause an artificial appearance in which people look like their skin is made of plastic. This represents the most widely used noise-reduction approach.

A **temporal filter** takes advantage of the fact that noise is a random element of the image that changes over time. Instead of simply evaluating individual frames, a temporal noise filter evaluates several frames at once. By identifying the differences between two frames and then removing that data from the final image, visible noise can be reduced very effectively. If there are no objects in motion, this is a virtually perfect noise-reduction technique that preserves most of the detail. This approach is used by many high-end competitors.

However, a problem arises if there are moving objects in the image, which also cause differences from one frame to the next; of course, these differences should be retained. If moving objects are not distinguished from noise, a ghosting or smearing effect is seen.

HQV processing uses a per-pixel motion-adaptive and noise-adaptive temporal filter to avoid the artificial appearance and artifacts associated with conventional noise filters. To preserve maximum detail, moving pixels do not undergo unnecessary noise processing. In static areas, the strength of noise reduction is determined on a per-pixel basis, depending on the level of noise in the surrounding pixels as well as in previous frames, allowing the filter to adapt to the amount of noise in the image at any given time. The end

result is a natural-looking picture with minimal noise and grain and maximum preservation of fine details.

## Codec Noise Reduction

A second type of noise common to digital cable or satellite or shows recorded with a PVR such as a TIVO is mosquito noise. While random noise is random, mosquito noise has a specific pattern. It is the white frizzy and blotchy distortion that is seen at the edges of objects, often appearing like mosquitoes buzzing around that object. "Block noise" is a third type of distortion in which artificial horizontal or vertical lines produce the appearance of blocky edges. Both of these types of noise are caused by high levels of video compression used by broadcasters to fit more television channels within a given bandwidth, by personal video recorders to increase the maximum recording capacity. As more digital channels are added, this will become a bigger problem.

HQV video processors are able to individually identify these three types of distortions and separate them from details in the video. Thanks to HQV's true per-pixel processing, different approaches to noise reduction can be used for each pattern of distortion. The result is superior noise reduction that preserves more detail than what would normally be possible.

## Detail Enhancement

Detail enhancement, also called sharpening, is a necessary component of all digital imaging, both standard definition and high definition. Unfortunately, due to the historically poor implementations of sharpening algorithms, this process has received a reputation as something to avoid.

All digital video goes through a lowpass anti-aliasing filter to prevent false color and moiré effects that can occur during the digitization process. The filter improves overall image quality, but it necessarily blurs some of the detail. The data-compression stage can also remove some detail. Fortunately, much of the lost detail can be mathematically recovered.

Because the human visual system perceives sharpness in terms of apparent contrast, exaggerating the differences between light and dark can produce what appears to be a sharper image. Unfortunately, due to rudimentary implementations of sharpening in the past, this process has been associated with artifacts known as "ringing" or "halos" in which objects are surrounded by a bright white edge. The resulting image appears harsh and does not reflect what was originally captured. The halos can sometimes be more distracting than the softness from the uncorrected image. For that reason, it is often recommended that users turn down the sharpening on video devices.

HQV Detail Enhancement technology is different. By using a more conservative algorithm and selectively identifying the area of blur before processing, HQV Detail Enhancement avoids halo or ringing artifacts at even the highest setting. Of course, it is also possible to disable HQV Detail Enhancement if the source has already applied sharpening. A key benefit of HQV Detail Enhancement is that, when used in conjunction with our 1024-tap scaler, standard-definition TV can be delivered at near high-definition quality.

## **10-bit 4:4:4 Internal Data Paths**

Not only does HQV processing implement some of the most advanced algorithms for video processing, but the internal data paths support 10-bits-per-channel with full 4:4:4 color sampling. (The term "4:4:4" refers to the fact that the color information can be input at full horizontal resolution, and 10-bit data paths provide 1024 steps of brightness and color.) The result is the ability to render over 1 billion colors. In comparison, conventional video processors that only have 8-bit data paths result in the ability to render only 16 million colors. Simply put, by maintaining more bits in the data, HQV products can preserve all the fine detail and dynamic range found in the original source.

## **Summary**

As you can now understand, HQV processing represents an enormous leap in video processing, with true flagship performance in de-interlacing, noise reduction, and scaling with both standard-definition and high-definition signals. Silicon Optix designed HQV processing as a no-compromise solution.