WHITE PAPER

A Designer's Guide to HD Video Pre- and Post-Processing

Make your HD multimedia design stand out in crowded consumer markets with video pre- and post-processing

HD digital video is revolutionizing multimedia consumer products—from small, handheld personal media players and mobile phone handsets to giant wallscreens. HD video encoding and decoding algorithms get the bulk of attention by designers developing such products, but HD video codecs for advanced video applications such as broadcast television and Blu-ray disc players have quickly become standardized; there's little room for product differentiation in a standardized video codec. However, you can substantially differentiate an HD product's design by improving the video image stream—both before video compression and after.

Consumer video moves to the forefront once again

There's an explosion occurring in consumer video. It's fueled by the adoption of digital video-compression standards, which deliver excellent picture fidelity and give consumers the ability to shape and mold video for their own viewing preferences. The popularity of digital video has revived the consumer television market, boosting televisions from the low-margin, sub-$100, bargain-basement category to highly featured, wall-size flat screens that cost many thousands of dollars. At the same time, HD disc players; DVRs; cable, satellite, and terrestrial boradcast set-top boxes; and camcorders also lure their share of consumers' wallets and pocketbooks. Entirely new product categories are starting to appear, such as digital SLRs with integrated HD video recording that have been introduced by several of the top camera vendors. After decades of slumber, video is once again a hot commodity in consumer electronics.

Consequently, product differentiation through video-enhancing features has also become important. Consumers directly compare picture quality between products that are on display in stores and make purchasing decisions based on perceived picture quality, image fidelity, and color presentation. While digital video-compression codecs are at the heart of the video revival as a hot consumer-product category, these codecs are now standardized for HD (high-definition) video products and do not offer much room for product differentiation. Instead, consumer-product developers have discovered that video pre- and post-processing algorithms (performed before video encoding and after video decoding) are the tools they need to make their HD offerings stand out in a very crowded market.

Digital video history

Without doubt, digital compression technology has remade the world of consumer video. At the beginning of the 1990s, digital video was rare and analog video ruled. This was the VCR era. During the early part of the 1990s, the electronics industry dabbled with various digital video experiments such as the video CD (VCD), digital satellite broadcasting, and various video-streaming computer file formats. Many of these early experiments with digital video used the MPEG-1 video-compression standard, which was developed in the late 1980s and finally approved in late 1992. Perhaps the most notable legacy of the MPEG-1 video experiments was the development and wide adoption of the MP3 (MPEG-1 Layer III) digital-audio compression format, which took the world by storm and became the foundation of the personal digital-audio player.

However, the quality objective for MPEG-1 video—matching the picture fidelity of VHS videotape—set a fairly low bar. MPEG-1 video had obvious, visible image flaws and lacked support for interlaced video. As a result, MPEG-1 video compression was shortly replaced with the MPEG-2 video-compression standard, which was announced in 1995 and approved in July, 1996. The first consumer product to adopt MPEG-2 video-compression appeared just a few months later. It was the DVD and consumer players and discs became available in Japan by November, 1996. That's when Toshiba launched its SD3000 DVD player. DVD discs and players were on sale in the United States by March, 1997. With its vastly improved picture quality, MPEG-2 rapidly became the most widely used digital video-compression standard for SD (standard-definition) television. In addition to its use in DVD players, digital satellite broadcasters rapidly adopted MPEG-2. Cable operators followed soon after. Meanwhile, a number of other video-compression standards for computer-based SD video have also appeared.

With the turn of the century and of the millennium, consumer video's focus has turned from SD to HDTV, which required an upgrade to the state of the art in video compression. The answer to this need is more advanced video-compression codecs such as the H.264 AVC video codec, which is rapidly being adopted across the board for HDTV as employed in Blu-ray disc players (the DVD replacement), HD cable and satellite broadcasting, video camcorders, and computer-based video. The happy result for product vendors: higher prices for all manner of consumer video products. Televisions, which cost one or two hundred dollars for standard-definition units with

CRT displays now sell for many hundreds or thousands of dollars for the larger flat-panel HD units. DVD players, which had dropped in price to as low as $20 now sell for $250 or more in their Blu-ray disc reincarnation. HD camcorders cost roughly twice what their SD counterparts sell for. With this new, higher price ceiling, there's much more room for additional features and better product differentiation.

Differentiating video products

HD digital-video codecs are standardized to ensure interoperability. Standard encoders must be able to drive standard decoders or interoperability suffers, as will sales for the entire category. Consequently, there's not much room for differentiation within the digital-video codecs. That's a potential disadvantage for vendors that wish to differentiate their products but it's an advantage for end-product designers because codec developers can design very efficient hardware to implement the codecs. The rapid standardization of HD codecs means that the codecs aren't likely to change much in the future (at least not for the HD product generation) and therefore need not incorporate much programmability.

The opportunity for product differentiation and the need for programmability now reside in the video pre- and post-processing blocks that improve upon the picture and color fidelity delivered by the digital HD codecs. Figure 1 presents a simple pair of block diagrams that show exactly where this opportunity for differentiation resides. In products that capture video, such as cameras and camcorders, the opportunity resides in pre-processing the video stream between the video-capture sensor and the video encoder. For display-oriented products such as Blu-ray players, set-top boxes, and HDTVs, the opportunity resides between the video decoder and the product's video output or display.

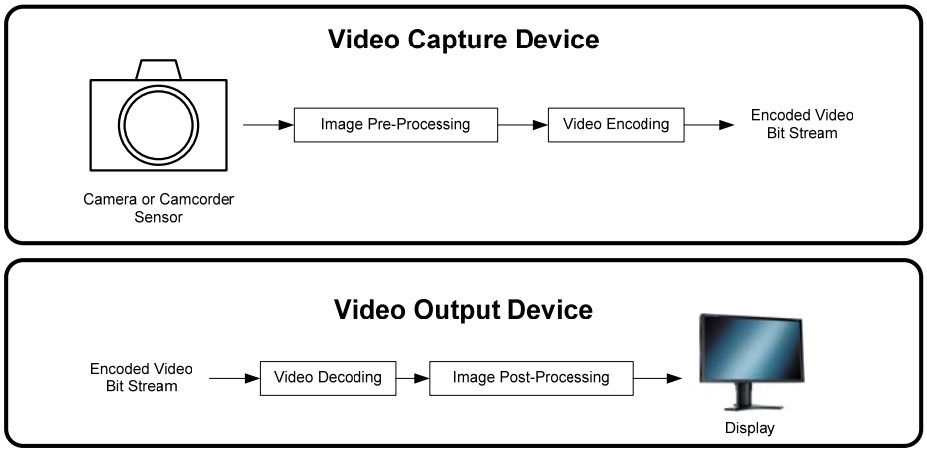

Figure 1: Basic block diagrams for video-capture and video-output devices showing where video pre-processing and post-processing occur.

Video Pre-Processing Algorithms

Unsurprisingly, video doesn't stream off of a sensor in pristine condition. There is a long list of transformations that can be performed on this raw video to improve the image before it's encoded. Some of the pre-processing operations include:

- Pixel scan/data transfer – This operation simply gets the image off of the sensor. Most image sensors have a variety of required control sequences for powering up and initializing the sensor, freezing an image in the sensor, and reading out the image. All of these tasks require the attention of an off-sensor controller.
- Bayer pattern deinterleaving – Modern video imaging places a 3-color RGB Bayer filter over a monochromatic image sensor. The data streaming off of the sensor therefore contains red, green, and blue information that must be separated so that the image can be converted to YCbCr luma and chroma representation.
- Noise filtering – Nothing electronic occurs in the absence of noise. The best place to filter or eliminate noise is before encoding. There's no need to spend effort encoding the high-frequency image variations attributable to noise.
- Shake detection and compensation – All cameras and camcorders shake, by varying amounts. The shaking blurs and degrades image quality. Shake detection and compensation can reduce the image degradation caused by camera shake.
- Localized dynamic range compensation – An image's dynamic range can exceed that of a sensor operated in a particular capture mode. An appropriately equipped camera or camcorder can intelligently adjust exposure within a frame to expand the sensor's basic dynamic range.
- Focus adjustment (sharpening) – Image sensors do not capture continuous images. Instead, they break the image into pixels, which must then be interpolated and reassembled into an image at a specified resolution. These operations cause a loss of sharpness, which can be corrected with the appropriate pre-processing algorithm. In addition, an image that's slightly out of focus because of an improperly focused lens can be brought into sharper apparent focus using a sharpening algorithm.
- Color correction – White light varies depending on color temperature, which affects the resulting captured video. In addition, different image-presentation systems alter the color in various ways. Video color correction compensates for these color-shifting factors, which are especially apparent with skin tones.
- Face detection – In images populated with people, the videographer generally wishes for the faces to be in focus. A video-capture device that can recognize faces can do a better job of setting the camera's or camcorder's focus.

- Stereoscopic imaging – It is possible to synthesize stereoscopic images from density information in a flat image.[1] This sort of pre-processing can be used, for example, in mixed-reality systems that combine virtual and real imagery. In addition, many future cameras will have true, 2-lens stereoscopic imaging. Both views will be displayed by a 3D display. Many Hollywood movies are already being "filmed" stereoscopically.

Figure 2 shows graphic examples of four pre-processing algorithms: noise filtering, shake detection and compensation, dynamic range compensation, and color correction.

[1] Hierarchical Depth Estimation for Image Synthesis in Mixed Reality, HanSung Kim and Kwanghoon Sohn, Dept. of Electrical and Electronic Engineering, 134 Shinchon-dong, Seodaemun-gu, Seoul, Korea 120-749, http://www.3dkim.com/Eng/2-04.pdf

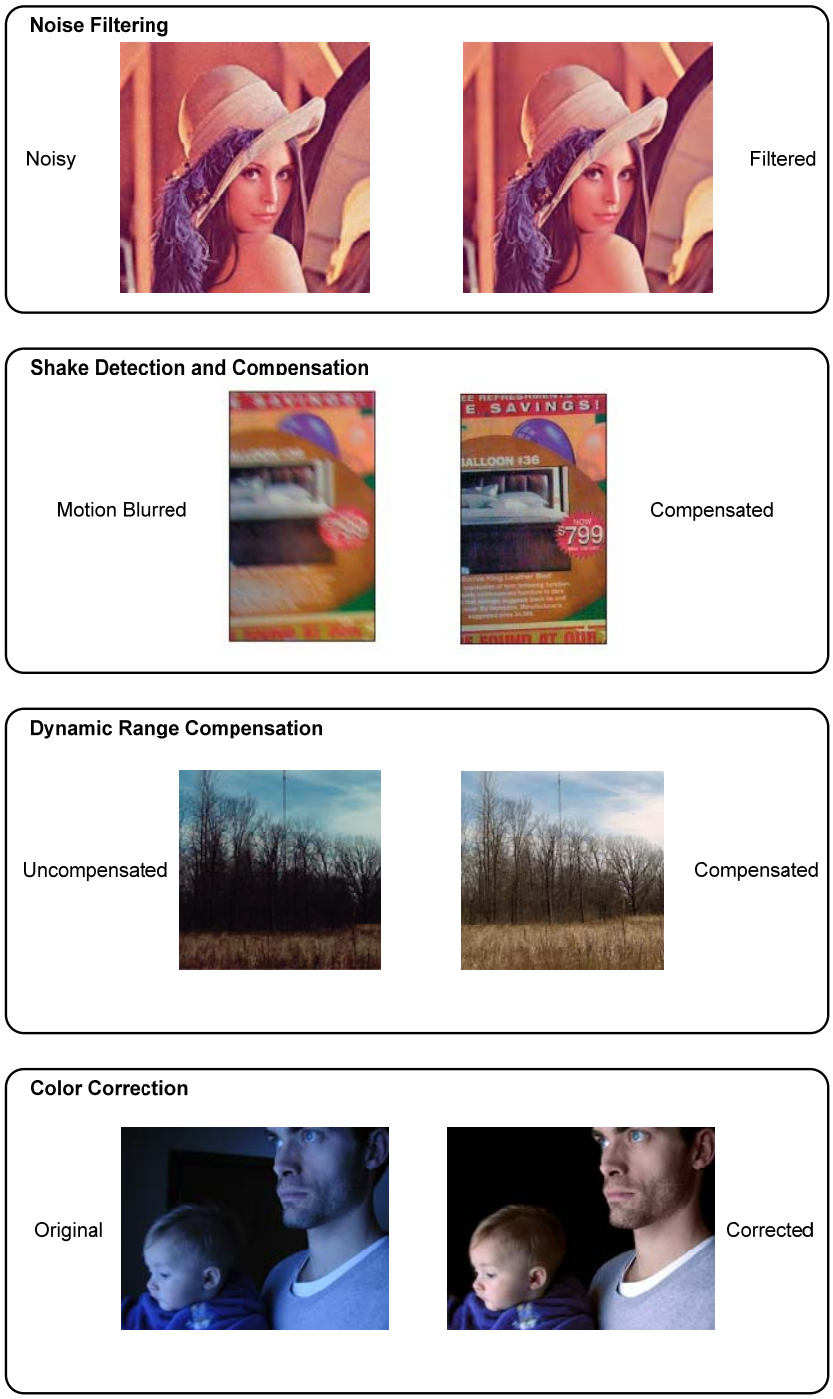

Figure 2: Examples of four pre-processing algorithms: noise filtering, shake detection and compensation, dynamic range compensation, and color correction

Video Post-Processing Algorithms

Video emitted from a standardized video decoder can also have many imperfections and post-processing algorithms can greatly enhance the resulting image. Products with visibly better images sell better for higher prices. Video post-processing operations include:

- Deblocking/deringing filters– Video compression and decompression break images into blocks, encode them, and then put the blocks back together at the other end. These operations leave visible artifacts that can be reduced through deblocking/deringing filters.
- Edge detection – Edge-detection improves scaling, deinterlacing, and other video-processing operations. In addition, edge detection is important in video applications such as security and video surveillance, traffic management, and medical imaging.
- Scaling – With the explosion of new wall, desktop, and handheld video devices, screen sizes vary widely. Chances are good that an encoded video stream will not be the correct size for many of these video devices. Image scaling solves this problem by resizing the image size to fit the screen.
- Deinterlacing - Interlaced video is designed for television viewing: each interlaced video frame is displayed as two consecutive fields, each field containing the odd- or even-numbered lines of video. Each field therefore contains half the lines of the frame in an alternating fashion. However, video displayed on computer and TV LCDs needs to be progressively scanned where every line of video appears in each consecutive frame. Interlaced video must be de-interlaced before display on a non-interlaced screen.
- Frame rate conversion – Television frame rates are standardized but can vary from country to country or application to application in the case of computer video. Video post processing can transform video from one frame rate to another. Video frame-rate upconversion through interpolation can increase the perceived smoothness of motion in displayed video by creating interpolated video frames between the decoded frames emerging from the video decoder. This ability can also be important for high-end, 120-fps displays and for upconverting 24-frames/sec film-based video to normal video rates.
- Noise filtering – All manner of noise sources can degrade video images. There are a variety of noise-filtering algorithms that improve the picture depending on the noise that's present.
- Video overlays/alpha blending – Many video systems generate imagery such as user interfaces that must be overlaid on top of live video. Alpha blending gives overlay images partial transparency.
- Color-space conversion/brightness/contrast/gamma correction – different displays have different dynamic ranges and handle colors differently (different color transfer functions). Appropriate video post processing can make the most of a display when its characteristics are known. Gamma correction is another important processing step because LCD panels each have unique transfer functions.

Figure 3 shows examples of four post-processing algorithms: deblocking and deringing, edge detection, image scaling, and deinterlacing.

Implementing Video Pre- and Post-Processing Algorithms

Unlike video compression/decompression codecs, video pre- and post-processing algorithms are not standardized. They vary by end product. In addition, these algorithms are constantly being updated and new ones regularly appear. As a result, hardware that implements video pre- and post-processing algorithms needs far more programmability than does video codec hardware. Processors that are specially configured to run these video pre- and post-processing algorithms are therefore a good design solution for such tasks.

You may not be familiar with the concept of customizable processors but you are no doubt familiar with standard, general-purpose processor cores, which have been in use since the first days of SOCs. These processor cores extend the long tradition of using standard processor architectures, first established back in 1971 with Intel's introduction of the 4004, the industry's first commercial single-chip microprocessor. Over the next several decades, system designers became accustomed to using processor architectures developed by others. They adapted these general-purpose processor cores as best they could and added acceleration hardware when the processor lacked the performance to handle all the required tasks.

The advent of SOCs—which are nothing more than ASICs with on-chip processors—in the mid 1990s changed the equation with respect to microprocessors. As long as the silicon is going to be custom tailored for a specific application, the on-chip processors can be custom tailored as well. However, the relative lack of processor designers and the complexity of developing and maintaining an associated tool chain have stopped designers from using custom-tailored processors in the past.

Automated tools are now available that allow logic designers and software developers—specifically not processor designers—to custom tailor processors for specific on-chip SOC tasks. The customized processors are optimized to run targeted algorithms faster than can general-purpose processors and DSPs because new registers and instructions have been added so that multi-instruction sequences become single-instruction sequences. Because they are guaranteed correct-by-construction, these processors require much less gate-level verification. Because all of the needed processing capability comes in one optimized IP core, design is simplified.

It should not be surprising that it's possible to develop video-specific processor architectures that deliver substantial performance boosts to various pre- and post-processing video algorithms. Let's look at some ways of tackling some video algorithms. The first way is simple: take the algorithm written in C, compile it for a processor, run it, and profile it. It's very easy to see the performance difference between processor architectures using this approach.

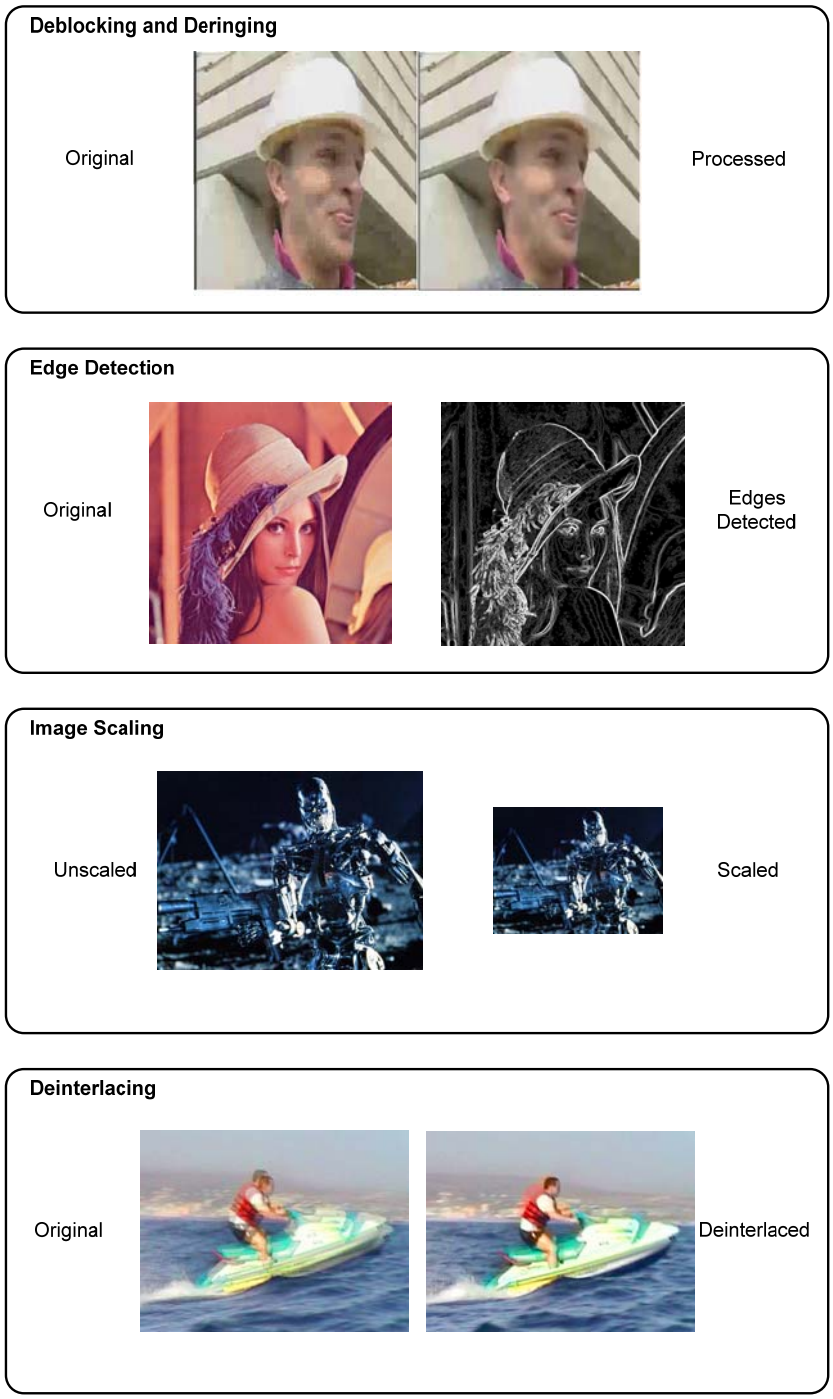

Figure 3: Examples of four post-processing algorithms: deblocking and deringing, edge detection, image scaling, and deinterlacing

Table 1 shows the result of such an approach for five different video algorithms running on a RISC processor and a DSP by comparing the required clock rates for both processors to execute each algorithm on the 62-Mpixel/sec image stream for 1080p video. The RISC processor in this example is a version of Tensilica's Xtensa RISC core. As you can see from the table, a 32-bit RISC processor would need to run at many GHz to execute these algorithms. Clearly, that's not possible in any currently available ASIC process technologies, which currently support processor clock rates of a GHz or less.

The second column in Table 1 shows the required clock rates for a vector DSP running the same compiled algorithms. In fact, the target DSP is also a version of Tensilica's RISC core, but the processor core has been extended with a comprehensive set of vector DSP instructions including a 4-way SIMD MAC and DSP ALU extensions that can perform 4- and 8-way SIMD vector operations. The processor has also been augmented so that it can issue three independent instructions per clock cycle. Collectively, the instruction extensions that transform the RISC processor into a vector DSP are called "Vectra LX." Even though the two processors shown in Table 1 are quite different, they are both built on the same architectural foundation and are programmed using the same set of software tools.

Table 1: Video algorithm performance results: RISC processor vs. DSP

Benchmark	Results	
	Pure C	Vectra LX
5x5 2D Filter	10.1 GHz	440 MHz
5x5 Threshold Filter	20.96 GHz	459 MHz
Median Filter	2.54 GHz	118 MHz
Frame Subsample	192 MHz	19 MHz
Frame Sum	273 MHz	50 MHz
Area	0.25 mm^2	0.63 mm^2

Table 1 notes:

1. Area values are for TSMC 65LP process, synthesized to 300MHz target

The video-processing algorithms[2] listed in Table 1 include:

- 5x5-Pixel 2D Filter: Two-dimensional filters can be used for several operations including sharpening images, adding blur, or to add dithering—which injects noise into the image to prevent the loss of detail caused by quantization and by restricted color palettes.
- 5x5-Pixel Threshold Filter: Threshold filtering reduces image pixel values to one of two values—Zero (whiter than white) or One (blacker than black)—effectively creating 1-bit pixels. Pixels with a starting value below the threshold value are set to 0 and pixels with a starting value above the threshold are set to 1. The result is a strictly black-and-white image with no shades of gray and the resulting image is used, for example, by noise-detection algorithms.
- Median Filter: A noise-reduction filter that replaces a pixel value with the median value of the surrounding neighbors. This sort of filter is often used to reduce the effects of salt-and-pepper or Gaussian noise.
- Frame Subsample: A simplistic image downscaler.
- Frame Sum: Adds the pixel values of two or more video frames. This algorithm can be used to measure the amount of light in different portions of the video frame, which is used for noise reduction, exposure compensation, and to aid some types of video compression.

Unsurprisingly, the SIMD/DSP-augmented version of the Xtensa processor is much more efficient than the unaugmented version and therefore requires a much lower clock rate to execute these video-processing algorithms for a 1080p video stream (or any other video stream for that matter). The performance improvement is about an order of magnitude or more (sometimes much more) for most of the video-processing algorithms and the required clock rates for the vector DSP are actually achievable using available ASIC manufacturing processes, as opposed to the unattainable multi-GHz clock rates for the 32-bit RISC processor. Yet the Vectra LX extensions increase silicon area very little in absolute terms (about 0.4 mm2 for TSMC's 65LP process).

Table 1 illustrates the performance boost that a good vector DSP can give to video processing. Table 2 shows what can be done when the processor is specifically customized for the target algorithms.

[2] Fundamentals of Image Processing, Hany Farid, http://www.cs.dartmouth.edu/farid/tutorials/fip.pdf

Table 2: Video algorithm performance results: RISC processor vs. DSP vs. processor customized with Tensilica's XPRES compiler

Benchmark	Results		
	Pure C	Vectra LX	XPRES
5x5 2D Filter	10.1 GHz	440 MHz	291 MHz
5x5 Threshold Filter	20.96 GHz	459 MHz	440 MHz
Median Filter	2.54 GHz	118 MHz	50 MHz
Frame Subsample	192 MHz	19 MHz	12 MHz
Frame Sum	273 MHz	50 MHz	25 MHz
Area	0.25 mm^2	0.63 mm^2	0.59 mm^2

Table 2 notes:

1. Area values are for TSMC 65LP process, synthesized to 300MHz target

The first two results columns in Table 2 merely repeat data from Table 1 while the rightmost column provides performance numbers in terms of required clock rate for a version of Tensilica's Xtensa processor that has been customized to efficiently handle all five video algorithms. The XPRES compiler accepts C source code and automatically generates processor enhancements that accelerate performance for the submitted code. The resulting processor is slightly smaller than the Xtensa processor core with the VectraLX extensions yet it outperforms the VectraLX version of the Xtensa processor by a factor of two or more on most of the algorithms. The required clock rates for the algorithms are actually quite low, which will result in substantial power and energy savings. The 5x5 threshold filter is an exception and both the VectraLX-enhanced and XPRES-enhanced processors exhibit similar performance for this algorithm.

Table 3 shows the results for the 5x5 2D and threshold filters and the frame summation algorithms when the processor has been directly customized for three of the algorithms.

Table 3: Video algorithm performance results: RISC processor vs. DSP vs. processor customized with Tensilica's XPRES compiler

Benchmark	Results			
	Pure C	Vectra LX	XPRES	Direct Customization
5x5 2D Filter	10.1 GHz	440 MHz	291 MHz	60 MHz
5x5 Threshold Filter	20.96 GHz	459 MHz	440 MHz	93 MHz
Median Filter	2.54 GHz	118 MHz	50 MHz	
Frame Subsample	192 MHz	19 MHz	12 MHz	
Frame Sum	273 MHz	50 MHz	25 MHz	< 6 MHz
Area	0.25 mm^2	0.63 mm^2	0.59 mm^2	0.70 mm^2 (See Notes)

Table 3 notes:

1. Area values are for TSMC 65LP process, synthesized to 300MHz target
2. 2D Filter area increment 0.35 mm2 over base core
3. Threshold Filter area increment 0.09 mm2
4. Frame Sum area increment 0.01 mm2
5. All area values are for TSMC65 LP, 300MHz target

Again, the first three result columns in Table 3 repeat data from the previous table. The last result column shows the result of adding custom instructions and enhanced I/O interfaces to the processor core for these particular tasks. The required clock rates are substantially lower for the three optimized algorithms and all require less than 100 MHz of processor performance. These results are quite substantial and are possible only by customizing the processor core to the task at hand.

Conclusion

Product differentiation through video-enhancing features has become critically important in consumer video markets. Consumers directly compare picture images and choose purchases based on the best perceived picture. While digital video-compression codecs are at the heart of the video revival as a hot consumer-product category, these codecs are standardized and do not offer much room for product differentiation. Consequently, video pre- and post-processing algorithms can help these video products stand out in a very crowded market and customized processor cores provide a fast, easy path to implement these algorithms with all of the performance your project requires.

An Open Invitation

To start your design team on that path to the successful design of a video-processing SOC, contact Tensilica for assistance. For more information on the unique abilities and features of the customizable family of Xtensa processor cores and associated development tools, see www.tensilica.com, send an email to sales@tensilica.com, or contact Tensilica directly at:

US Sales Offices:

Santa Clara, CA office:
3255-6 Scott Blvd.
Santa Clara, CA 95054
Tel: 408-986-8000
Fax: 408-986-8919

San Diego, CA office:
1902 Wright Place, Suite 200
Carlsbad, CA 92008
Tel: 760-918-5654
Fax: 760-918-5505

Boston, MA office:
25 Mall Road, Suite 300
Burlington, MA 01803
Tel: 781-238-6702 x8352
Fax: 781-820-7128

Austin, TX office:
Tel: 512-413-4410

International Sales Offices:

Yokohama office (Japan):
Xte Shin-Yokohama Building 2F
3-12-4, Shin-Yokohama, Kohoku-ku,
Yokohama
222-0033, Japan
Tel: 045-477-3373 (+81-45-477-3373)
Fax: 045-477-3375 (+81-45-477-3375)

UK office (Europe HQ):
Asmec Centre
Eagle House
The Ring
Bracknell
Berkshire
RG12 1HB
Tel : +44 1344 38 20 41
Fax : +44 1344 30 31 92

Israel:
Amos Technologies
Moshe Stein
moshe@amost.co.il

Beijing office (China HQ):
Room 1109, B Building, Bo Tai Guo Ji,
122th Building of Nan Hu Dong Yuan, Wang Jing,
Chao Yang District, Beijing, PRC
Postcode: 100102
Tel: (86)-10-84714323
Fax: (86)-10-84724103

Taiwan office:
7F-6, No. 16, JiHe Road, ShihLin Dist,
Taipei 111, Taiwan ROC
Tel: 886-2-2772-2269
Fax: 886-2-66104328

Seoul, Korea office:
27th FL., Korea World Trade Center,
159-1, Samsung-dong, Kangnam-gu,
Seoul 135-729, Korea
Tel: 82-2-6007-2745
Fax: 82-2-6007-2746