

Tensilica White Paper

Frame Processing Time Deviations in Video Processors

May, 2008

Executive Summary

Chips are increasingly made with processor designs licensed as semiconductor IP (intellectual property). Video processing is no exception. The nature of video coding is such that some frames of video take longer to process than others. Wide variations in processing time make correct operation of the chip and the final system unpredictable. A video processor that minimizes the variations in processing time for each frame enables a more reliable, less expensive, and lower power system design. The Tensilica Diamond 388VDO processor excels at minimizing deviations in frame processing time.

How Video Is Compressed

Whether video is processed in a hardware block, a general purpose processor, or an optimized DSP processor, each frame takes a different amount of time to process. This is because each frame of video is different and is compressed in variable ways by the encoder for best efficiency.

Most video coding standards process video as a sequence of square macroblocks. One important technique for video compression is identifying and coding macroblocks in each video frame image that are identical or similar to their neighbor. For example, the sky at the top left corner of the image in Figure 1 is almost identical from one macroblock to its neighbor. Another important technique for achieving compression is by identifying objects that have moved from a nearby location in a previous frame of video, such as with the light post on the bridge.



Figure 1: An example frame showing the Vincent Thomas Bridge between Long Beach and San Pedro, California.

These techniques are known as intra-frame and inter-frame prediction, respectively. By using prediction, a video encoder must encode the location – in the current or a previous frame – from which to predict each macroblock and the imperfections between the prediction and the actual captured pixels. Fortunately the imperfections, known as residuals, tend to be much smaller data values than the actual captured pixels and so less data must be encoded than would be the case without prediction.

Naturally, video frames with a lot of macroblocks with fine details will be less accurately predicted from their neighbors and video frames with a lot of moving objects will be less accurately predicted from previous video frames. Therefore, the marathon frame in Figure 2(a) will require more data for prediction and yield less accurate prediction and therefore more data to encode the imperfections than the sky frame of Figure 2(b).

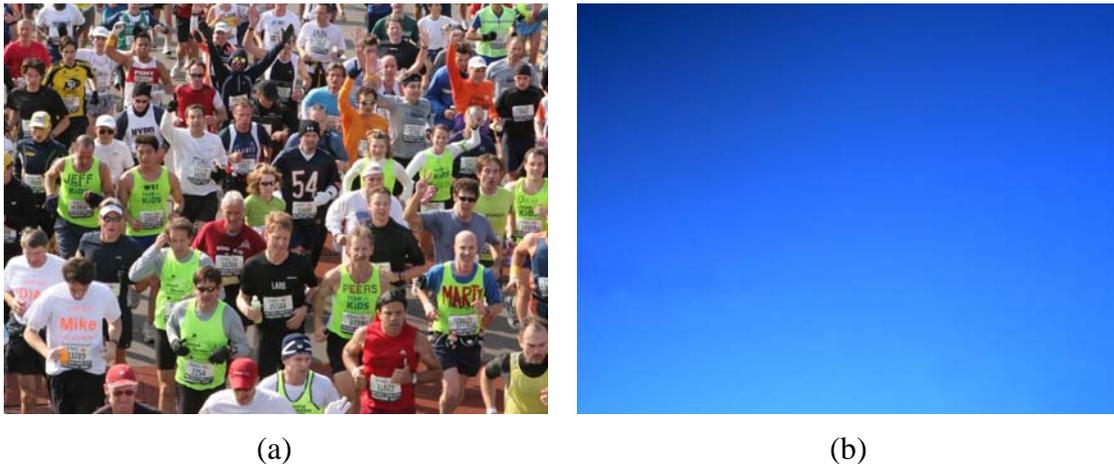


Figure 2: A frame from the 2006 New York City marathon and a frame of clear sky.

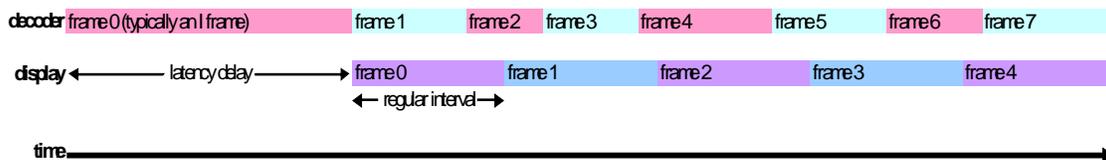
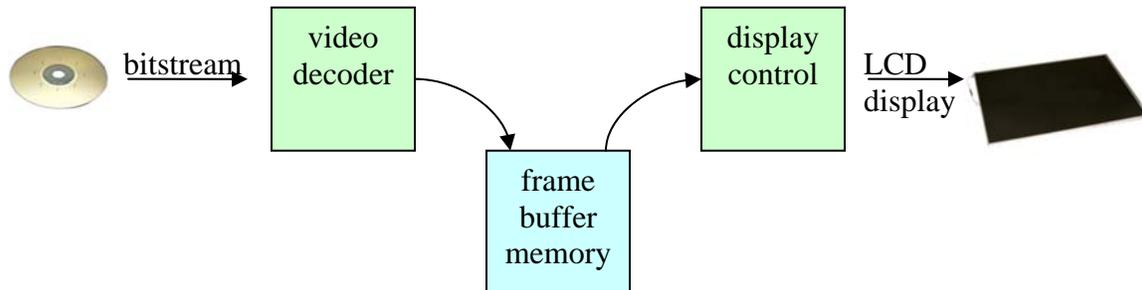
The Implications of Prediction Types

In video coding, some video frames are entirely intra-predicted. This avoids dependencies on previous frames, which ends the propagation from one frame to the next of any random or imprecision errors that accumulate. Because fully intra-predicted video frames rely only on prediction from neighbors in the same frame, the opportunity to make more accurate predictions from previous frames is lost. As a result, the predictions in intra-predicted frames are less accurate, the residuals values are greater, and the bitstream data for the compressed frame is greater than for inter-predicted video frames. Intra-predicted frames are typically coded with two or more times as many bits as inter-coded frames.

Because of the difference in the amount of data required to encode frames, almost invariably codecs take a longer time on intra-predicted (I) frames than on inter-predicted (P) or bi-inter-predicted (B) frames. Some video codecs take dramatically longer to process I frames while others exhibit more invariant processing times for the different frame types.

Video Device Design

Video entertainment devices such as Blu-Ray/DVD players, set-top boxes, and portable media players are designed to buffer several frames of video before they are displayed. By doing so, while the video processor decodes a difficult video frame, the output interface is still able to display video frames from the buffer at the correct frame rate. The buffer is then refilled as the decoder finishes easy frames early. A video processor with little deviation in processing time for different frames allows the video chip and system to be designed with less frame buffer memory and therefore at lower cost.



In real-time video systems, such as for 2-way conferencing or automotive cameras, processing latency is critical. Such systems can not tolerate the display latency introduced by buffering multiple frames and must run the video codec at a high enough clock speed to handle the worst possible video frame. A video processor with little deviation in processing time for different frames allows the video chip and system to be designed with a lower clock speed and therefore at lower cost, lower power consumption, and higher reliability.

Many chip makers selecting their first video codec IP core do not think to ask about deviation in frame processing time. Few IP core vendors care to measure or specify it. However, chip makers with deep expertise in video processing always ask this important question and use it as a key decision criterion in selecting their video codec core.

Calculating the Effect of Frame Processing Time Deviation on Clock Rate

The frame processing time for many video processors varies with significant deviation between intra- and inter-predicted frames. In Figure 3, for example, it is easy to identify the regularly occurring I frames by their spikes in processing time. The higher the spikes stand out above the average, the greater accommodations must be made in the system design to accommodate the difficult frames.

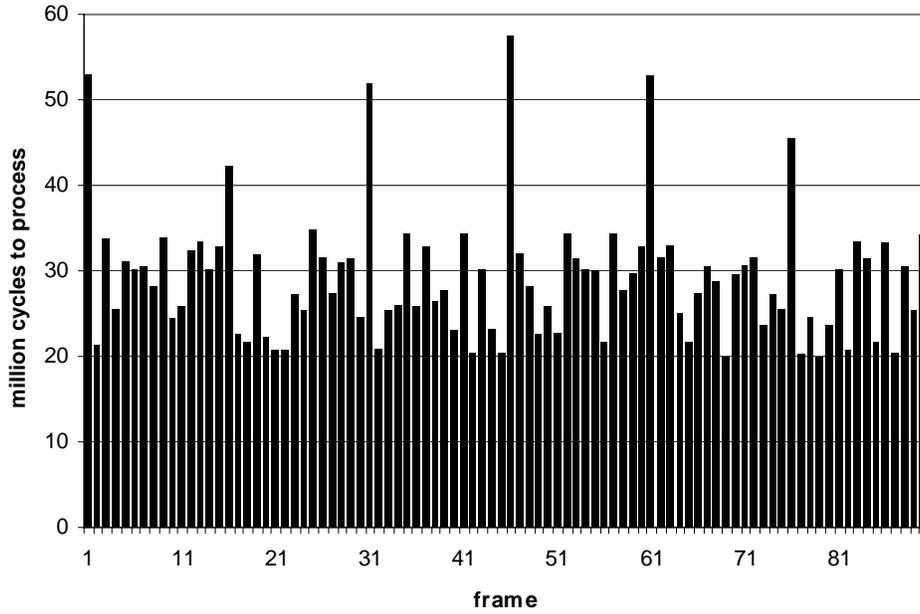


Figure 3: Video frame processing time on a processor with typical deviation.

The clock speed needed to perform video processing, F , is calculated as

$$F = R * \max_{f=0}^{N-B} \left(\frac{\sum_{i=f}^{f+B-1} P_i}{B} \right)$$

where R is the frame rate, N is the total number of frames processed, B is the number of display output frame buffers, and P_i is the processing time in cycles for frame i . This is the clock rate required to process the worst case window of video frames where the size of the window is the number of the display output frame buffers. Because I frames are generally a small portion of frames in video sequences and are almost invariably surrounded by faster-to-process P and B frames, this means that the frequency required to process video decreases dramatically with just an output display buffer that is just two frames deep.

In the case of display-latency-critical video processors such as for video conferencing and safety-critical applications, display frame buffering can not be tolerated. As a result, $B = 1$ and the frequency required is simply that for the worst case frame of video.

$$F = R * \max_{f=0}^{N-1} (P_f)$$

It can also be derived from the two equations above that clock rate required for processing a sequence of video frames is lower if the worst case frame requires less processing time.

Frame Processing Time Deviation for Diamond 388VDO

The Tensilica Diamond 388VDO video DSP processor and the software codecs that run on it have remarkably low deviation in frame processing time. Figure 4 shows an example of frame processing times possible with the processor for decoding H.264 video. By comparison to Figure 3, the deviation in frame processing time is lower and the worst case spikes in frame processing time are lower for Diamond 388VDO than for typical video processors.

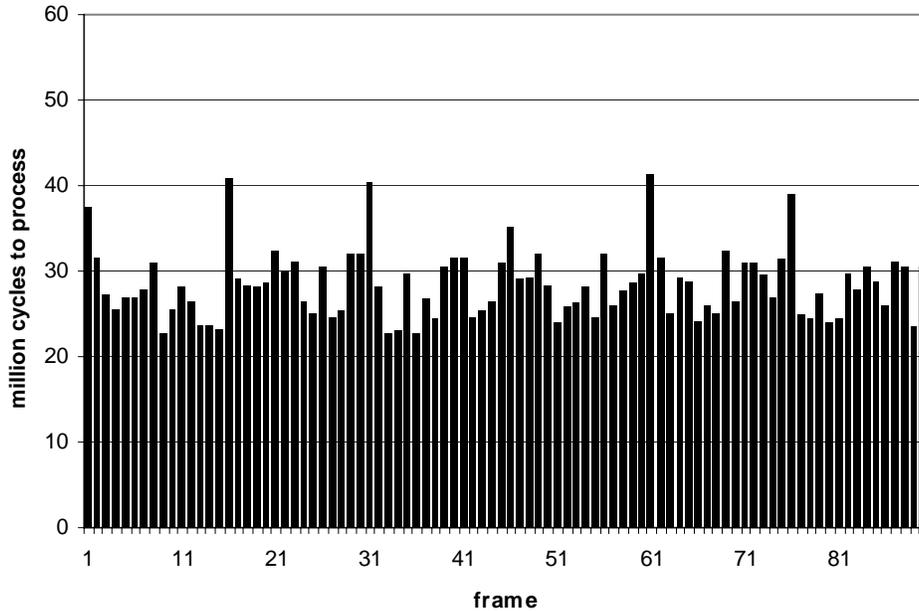


Figure 4: Video frame processing time on Tensilica Diamond 388VDO

A H.264 Baseline Profile video stream of movie trailer video sequence with a ratio of 60 P frames per I frame decodes on Diamond 388VDO with a standard deviation of only 22% of the average frame processing time. This pegs Tensilica’s Diamond 388VDO as a truly steady video codec processor.

Key Underlying Design Details

Diamond 388VDO achieves this reliability as a benefit of using the Tensilica Xtensa processor and Tensilica’s industry-leading tools for simulation, code profiling, and instruction set development. Xtensa is a uniquely configurable embedded processor architecture allowing highly application-specific performance optimizations. Determining the optimal configuration of an Xtensa processor is possible because of the fast and accurate processor simulations, the detailed code profiling, and the clear illustration from these tools of performance bottlenecks and ways that they can be removed with appropriate extensions to the processor. See the [Xtensa processor development tool kit product brief](#) for more information.

The Diamond 388VDO video processor is actually built from two Xtensa cores and a specialized DMA controller as shown in Figure 5. The separation of data processing into

two processors allows one, the Stream core, to manage the system and handle the compressed bitstream while the other, the Pixel core, handles the heavy duty DSP processing functions simultaneously. The DMA block moves data between external frame buffer memory and the internal scratchpad memories of the two cores. Efficient use of the DMA controller by the Stream core for prefetching allows the Pixel core to remain busy, avoiding processor stalls and giving comparatively invariant frame processing time.

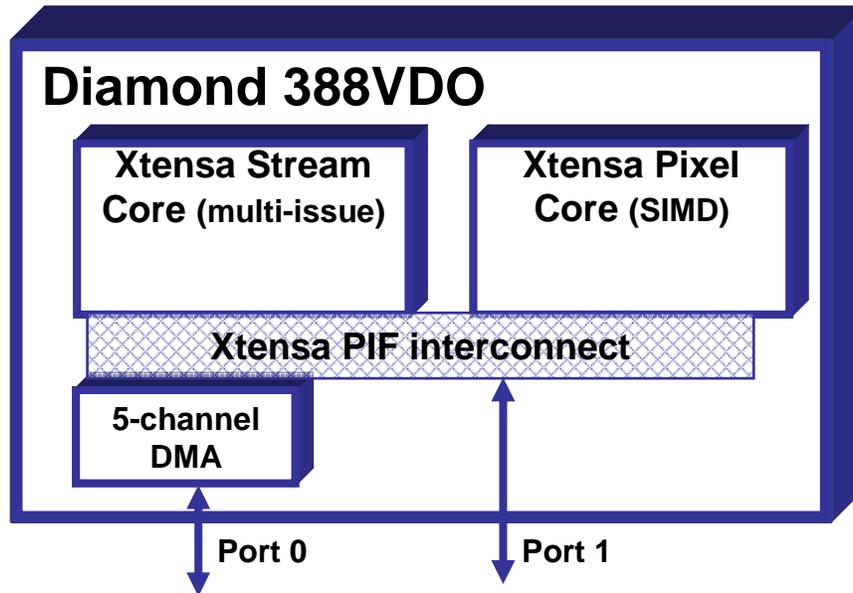


Figure 5: Block diagram of Diamond 388VDO

Because I frames achieve less compression than P frames they require more residuals data to be decoded from the bitstream. For video processors with separate stream and SIMD cores, this extra entropy decoding on the stream core makes it limit the overall throughput rate of the processor on I frames. This accounts for the amount of processing time in the I frame spikes of most processors. The instruction set extensions in the Diamond 388VDO stream core to accelerate entropy decoding are why the processor is less susceptible to long processing time of I frames than other video processors.

Summary

The design of Diamond 388VDO and its codec software to keep frame processing time deviation low make it dependable for many chip designs for which reliable performance is critical. This is a key factor in the decision of leading mobile entertainment and real-time video SOC designers to choose Diamond 388VDO as the video processor for their chips.

Tensilica, Inc.
3255-6 Scott Blvd.
Santa Clara, CA 95054

Phone: (408) 986-8000
FAX: (408) 986-8919