OL_H264e
HDTV H.264/AVC Baseline Video Encoder

Rev 1.0

General Description

The OL_H264e core is a hardware implementation of the H.264 baseline video compression algorithm.
The core accepts up to the highest resolution HDTV video stream as input and outputs the encoded bitstream.
Simple, fully synchronous design with low gate count.

Applications

- Digital video recorders.
- Video wireless devices.
- Video surveillance systems.
- Hand held HDTV video cameras.

Features

- Fully compatible with the ITU-T H.264 baseline specification.
- Proven in FPGA : VGA (640x480) at 30 fps in VirtexII-4 demo board.
- Profile level 4.1, can be decoded by Main Profile decoder.
- Supports up to the highest HDTV video resolution (1920x1080 @ 30 fps progressive).
- Very low operational frequency : from ~1.5 MHz for QCIF @ 15 fps to ~250 MHz for 1920x1080 @ 30 fps.
- Single core HDTV support in FPGA : 720p (1280x720) at 30 fps in high end FPGAs (Virtex4) . 4 CIF (704x576) at 30 fps in low end FPGAs (Spartan3-4, slowest speed grade).
- No CPU required for encoding.
- Variable Bit Rate (VBR) and Constant Bit Rate (CBR).
- Very low latency in VBR (~1.1 ms for VGA @ 30 fps).
- Motion vector up to –16.00/+15.75 pixels (search area is 32x32 pixel wide down to quarter pixel).
- Support for most of intra4x4 and all intra16x16 modes.
- Multiple slices support for better error resilience.
- Block skipping logic for lower bitrate.
- Deblocking filter for better quality.
- External memory interface tolerant of high latencies and delays, ideal in a SoC system or in a shared bus with a CPU. The memory interface can be clocked at a different frequency from the ocre for easier integration.
- Supports YUV 4:2:0 video input.
- Min Clock speed = 4 x the raw pixel clock speed.
- Low gate count : from 129K gates + 100 Kbits of RAM for real time VGA encoding to 178 Kgates + 100 Kbits of RAM for real time 1080p encoding.
- Simple, fully synchronous design.
- Available as fully functional and synthesizable VHDL or Verilog soft-core.

Functional Description

The OL_H264e core is a hardware implementation of the H.264 baseline video compression algorithm designed to process HDTV progressive video up to 1920x1080 at 30 fps.

Each block of 16x16 pixels is processed in just 1024 cycles. This means that each pixel is processed in just 4 cycles. Consequently, given an uncompressed video stream of resolution X by Y, and frame rate fps, the minimum clock frequency to process a such video stream is :

F = 4*X*Y*fps

This allows the core to process the video stream at relatively low clock frequencies. For example, HDTV video of 1920x1080 @ 30 fps requires ~250 MHz, whereas VGA video of 640x480 @ 30 fps requires ~37 MHz.

The table below summarizes the relationship between some possible video resolutions and frame rates and the clock frequency of the core.

Resolution	QCIF @ 15 fps	CIF @ 30fps	VGA @ 30fps	1280x720 @ 30fps	1920x1080 @ 30fps
Core freq.	~1.5 MHz	~12.1 MHz	~36.8 MHz	~110.5 MHz	~250.6 MHz

Table 1 Core frequency versus video resolution and frame rate.

A block diagram of the core is shown in Figure 1.

For each block, the intra prediction unit generates a suitable prediction. The intra prediction unit supports most of intra4x4 and intra16x16 modes.

In case of P-frames, the motion estimation unit generates a prediction as well. It examines an area of 32x32 pixels down to the quarter pixel (motion vector from –16.00 to +15.75). Quarter pixel prediction is generated using the tap filters described in the ITU-T specification.

The prediction of each unit is costed using Lagrange multipliers and the best is selected for encoding.

The residual information is calculated from the difference between the current block and the prediction. It is then transformed and quantized to be encoded by the lossless encoding unit.

The transformed, quantized residual is also used to reconstruct a reference frame, which will be used during the encoding of future P-frames. This is achieved by inverse quantization and transform of the residual, that is then added back to the prediction.

Finally, the reconstructed frame is filtered before being stored back in the memory.

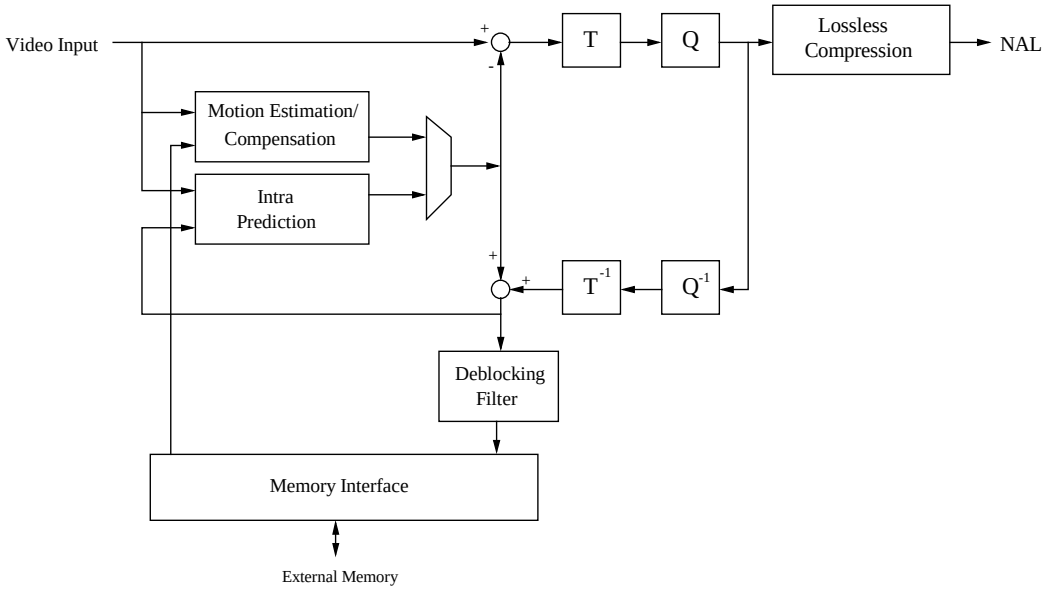

Figure 1 The OL_H264e block diagram.

Advantages of the core

Some of the key advantages of the core are discussed further below:

HDTV support

The core is designed to support up to the highest HDTV resolution, 1920x1080 @ 30 fps progressive. This opens a whole new range of applications from high-end video camcorders to high-resolution video surveillance at very low cost.

Low gate count

As it can be seen in the **Performance** section below, the core is designed very efficiently, with a low gate count. This allows 4CIF (704x576) video @ 30 fps to be processed by low end FPGAs at the slowest speed grade as well as 720p (1280x720) video @ 30 fps in high end FPGAs. Thus HDTV 720p real time encoding is possible in FPGAs without multiple core instantiation.

No external CPU required

The core can encode video independently, without the support of an additional CPU. This represents a large cost saving compared to solutions that require an external CPU.

Both VBR and CBR supported

The core supports both VBR (Variable Bit Rate) and CBR (Constant Bit Rate). This allows maximum flexibility for the designer.

Flexible memory interface

The core requires access to an external memory via a 32-bit data bus. About 50% of the whole theoretical bandwidth of the memory is actually used by the core.

This interface is designed to be independent from the memory used (i.e. DDR, SDRAM, SRAM, etc.).
More importantly, the memory interface is designed to tolerate high and unpredictable latencies and delays that are typical of a shared memory (i.e. AMBA and/or SoC where the bus is shared with a CPU or other cores).
In addition to this the memory interface can run at a different clock speed from the rest of the core. This simplifies the integration process and can save gates by not forcing the core to be synthetised to a much higher frequency just to be synchronous with the local bus.
This allows, for example, a core running at 37 MHz (encoding VGA @ 30 fps) to be easily integrated in a SoC sharing a 200 MHz bus with a processor.

Error resilience

The core supports multiple slices. This is useful in environments prone to data transmission problems (i.e. mobile phone or other wireless applications) in order to limit the damage inflicted to the image by transmission errors.

Low data rate features

The core supports two important features for low data rates: deblocking filter and macroblock skipping.
The deblocking filter especially improves the visual quality of the decoded image at low bitrates where the high quantisation noise produces unappealing "blockiness" in an image.

Macroblock skipping greatly reduces the bitrate with minimal effect on the visual quality of the decoded image.

Performances

Performance figures of the OL_H264e core implemented with some particular technologies are shown in the table below.

Technology	Approx Area	Speed	Video Throughput
0.13 u LV 0.9V, 125 C	178 Kgates + 106 Kbits RAM Optimised for speed	~ 250 MHz	1920x1080 (1080p) @ 30 fps
0.18 u slow process	129 Kgates + 106 Kbits RAM Optimised for area	~50 MHz	4 CIF (704x576) @ 30 fps
StratixII	17511 ALUTs + 5 M512 + 51 M4K + 3 DSPs	~118 MHz	1280x720 (720p) @ 32 fps
Virtex4-12	10,500 slices + 3 multipliers + 33 RAM blocks	~110 MHz	1280x720 (720p) @ 30 fps
Spartan3-4	10,500 slices + 3 multipliers + 33 RAM blocks	~50 MHz	4 CIF (704x576) @ 30 fps

Table 2 Performance of the OL_H264e core.

The gate count above does not include CBR and deblocking filter. Final results will be published shortly and they are not expected to increase by more than 10 Kgates.

OL_H264e : HDTV H.264/AVC Baseline Video Encoder

Summary

The combination of low gate count, low operating frequency, and full HDTV resolution support makes this core an application-enabling technology.

The applications of this core range from low power wireless application at relatively low resolution such as mobile phones to HDTV handheld recorders and video surveillance cameras.

The very small area of this core also allows novel applications such as its direct integration on to a CMOS sensor. This would create an extremely compact intelligent sensor accepting light directly at its input and outputting an H.264 bitstream.

Deliverables

Synthesizable VHDL or Verilog RTL.
Bit accurate C model.
Complete HDL testbench.
Complete data sheet.

Ocean Logic Pty Ltd
PO BOX 768 - Manly NSW 1655 – Australia Fax: +61-2-90120979
E-Mail: contact@ocean-logic.com URL : http://www.ocean-logic.com/