streamingnetworks
digital video delivered

iRecord Pro: Personal Media Recorder

Ayub Khan, Ph. D.

July 2008

Portable multimedia players ("PMP") have become very popular in recent years. Apple with its iPods® series was able to capture about 70% share of the PMP market in 2007. Apple strengthened this position further with the launch of iPhone and iPod touch in 2007. Microsoft® introduced its Zune® series of media players in 2006-2007. The Sony Play Station® Portable (PSP®) is a portable gaming console and with robust video and audio playback capabilities. Sony also launched the video Walkman in 2007. Other notable media player manufacturers include Archos, Creative Labs, iRiver, SanDisk. Market research firm iSuppli Corp. forecasts that global PMP unit shipments will increase to 268.6 million in 2011, representing a CAGR of 5.5% from 216.9 million in 2007.

Historically consumers have purchased PMPs primarily for their audio playback functions, but this is changing as the cost of incorporating video playback continues to decline. Market research from iSuppli projects that video enabled portable media players will outsell audio only models by the end of 2008, and will account for two-thirds of all PMP sales by 2011. Sony Walkman® now supports video as well. MP3 has beaten the early proprietary audio formats to become the near universal audio format, while several PMPs such as iPod and PSP also support Advance Audio Coding ("AAC") format.

A migration from proprietary digital video formats such as RealVideo®, QuickTime® and Windows Media Video®[1] ("WMV") to open formats with broad industry support is underway. Each successive generation of digital video format has utilized increases in available processing power to increase compression and enhance playback on both small screen portable devices and large screen high definition displays. MPEG-2 which was developed in the early 1990s is most prevalent format as it is widely used for DVDs, personal video recorders, and some digital television delivery systems. Early PMPs supported either WMV or MPEG-4[2], including the DivX and XviD variations on MPEG-4. Most modern PMPs now have sufficient processing power to decode the state-of-the-art H.264/AVC[3] video format which offers a number of improvements over both MPEG-2 and MPEG-4 format, especially for portable devices. The iPod, Zune, PSP, Sony Walkman, and numerous other PMPs now support H.264/AVC which is rapidly becoming the industry de facto video format.

PMPs store digital media as files on flash memory or on an internal hard disk. The iPod is the current capacity winner with up to 32GB of flash memory and up to a 160GB hard drive. As the capacity of flash memory keeps doubling

[1] *In 2006 Microsoft suceeded in having WMV9 officially approved as SMPTE 421M, also known as VC-1.*

[2] *Specifically MPEG-4 Part 2.*

[3] *H.264/AVC is also known as MPEG-4 Part 10.*

and price per Gigabyte keeps falling most analysts anticipate a broad shift to flash memory for all but the most capacity hungry users. While most PMPs have a USB 2.0 interface port they often have a proprietary connector, and in some cases also use a proprietary communications protocol. An increasing number of PMPs now support Wi-Fi, although the ability to update media over the Wi-Fi connection is only now being supported by their manufacturers.

Today the majority PMPs have QVGA displays (320 x 240), but new players are offering higher resolution as LCD panels prices continue to fall and consumers increasingly expect meaningful video support. As the PMP market matures the inclusion of robust video support is seen by many manufacturers as the key to getting otherwise content consumers to upgrade to a new device. The PSP has a 480 x 272 display, the iPhone® and iPod® touch each have a 480 x 320 displays, and although the iPod Classic still has a QVGA display it does support external video playback at full VGA resolution (640 x 480).

With PMPs available in a wide range of form factors from multiple vendors, and audio and video playback increasingly available in high-end mobile phones, consumers have plenty of choice to pick the device that best meets their needs. These devices hold out the promise of listening to favorite songs, TV shows and movies wherever and whenever the user wants. Unfortunately the reality often comes up short, especially when it comes to actually getting media onto the device.

Most PMPs either ship with or use broadly available software that lets users rip CDs to their device such as Apple's iTunes® or Microsoft's Windows Media Player. For computer literate users with lots of CDs and time on their hands this works well. Video is not so easy. Users can pay to download a hit TV series or popular movie from one of the online stores, but this is expensive, does not work with all devices, requires a fast internet connection, and will likely include heavy handed Digital Rights Management ("DRM") that may prevent a variety of legitimate uses. For less popular movies, home videos, or the box of old VHS tapes in the corner, the only option is to purchase and install video capture hardware into your computer, move your VCR or other playback device next to your computer, and spend hours capturing, encoding or transcending audio and video to the required format, and then transferring it to the player. This is after the user has figured out the appropriate encoding format, codecs, frame rate, frame size, bit rate, and a variety of other settings required to have the video playback in an acceptable fashion for each specific PMP device.

For most users there is no simple and easy way of getting their favorite media onto their portable media players. An ideal solution would be an easy to

streamingnetworks
digital video delivered

use device that can directly record audio and video onto any of the popular PMPs. Once the recording is finished it should be immediately available for viewing with no additional steps, or the need to use a computer. Streaming Networks introduced such a recording device, called iRecord® at CES 2007.

In one touch…iRecord

iRecord Pro is an easy to use device that enables users to record TV shows, movies, home videos, music, video gaming sessions, and even live events directly to the popular PMPs, with a touch of a button (Patent Pending). There is no downloading and no computer to configure, making it easy to set up iRecord Pro as part of an existing stack of AV equipment. It automatically detects the make and model of the PMP attached to it and generates media in the right format for that device (Patent Pending). iRecord Pro supports 3G iPhone, iPhone, iPod touch, iPod, PSP, Walkman, Creative Zen, Sandisk Sansa, Toshiba GigaBeat, Nokia N95 and USB flash drives, USB hard drives and numerous other devices[4].

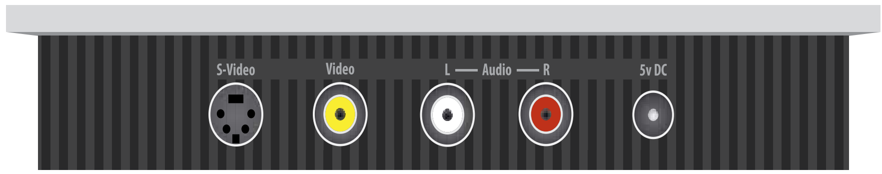
Figure 1: iRecord Pro

Source: Streaming Networks

The iRecord Pro is a compact device that includes all of the essential interfaces: a USB 2.0 host interface to connect the PMP or storage device, stereo audio line level inputs, a composite video and an S-video input. Almost any AV device manufactured anywhere in the world in the past 30 years can be connected using these standard inputs. iRecord Pro has just two buttons: one for power and the other to start and stop recording; LEDs indicate power and recording status. A remote control also provides the power on/off and record start/stop button as well buttons for pausing a recording and setting pre-defined duration recording of 30, 60, 120 & 180 minutes. However, the pause & duration recording functions can also be achieved using the power button of the unit. All the rest of the underlying complexity is handled automatically

[4] *See iRecord.com for a complete list of the devices currently supported by iRecord.*

iRecord Pro interfaces



Source: Streaming Networks

by iRecord Pro, including auto-switching between the video inputs; detection and correct processing of PAL, NTSC or SECAM signals, selection of 4:3 or 16:9 aspect ratio, de-interlacing, and numerous other settings.

For A/V recording video is digitally encoded in H.264 format and audio in AAC format at 128 kbps in stereo. Audio only recoding is in MP3 format at bit-rates ranging from 128 to 320 kbps in stereo. All MP3 players using Media Transfer Protocol (MTP) or mass storage protocols are supported.

Ease of use combined with outstanding video quality were the top priorities in developing iRecord. The temptation to add features and functions was ever present during design, but the desire to keep things simple kept winning the day. In the year since iRecord was launched user feedback has been overwhelmingly positive, especially as it relates to the ease of use. A frequent request has been for higher resolution encoding to support playback on new PMPs higher resolution screens, or external monitors connected to an iPod. As iRecord can record to most USB hard disks and flash drives, there were also a number of requests for higher resolution support from people watching video on laptops while traveling, or on a television hooked up to an AppleTV, Xbox360, or other devices that support H.264/AVC playback.

iRecord Pro, incorporates a number of exciting new features such as: H.264/AVC encoding at full D1 (720x480) and VGA resolution[5] (640x480) at an average bit-rate of 2.5 Mbps and 1.5 Mbps respectively; and the addition of a USB client port that enables recording directly to a PC or Mac computer hard drive. In addition to recording analog audio and video, iRecord Pro supports transcoding of MPEG2 digital video to H.264/AVC in hardware so users machines are not tied up for hours running software encoders.

Video Processing

The ease of use and simplicity of iRecord from a user's perspective belies the technical complexity inside. Consider video encoding: the H.264/AVC video standard (ISO/IEC 14496-10) represents the current state of the art video encoding designed to support both small file sizes for portable devices and high

[5] *Note that VGA is currently the highest resolution supported by any of Apple's iPod devices.*

definition video for larger displays. H.264/AVC has gained broad market acceptance and is a required format on both Blu-ray and HD-DVD players, and is implemented in iPod, iPhone, AppleTV, PSP, Walkman, Zune2, Xbox 360, and many other devices. The H.264/AVC standard delivers more than twice the compression of MPEG2 but at the cost of higher computational load: it is four to five times more computationally demanding than MPEG-2. This load increase is very significant in the consumer electronics domain where the success of a device depends largely on its cost competitiveness. Most consumer electronics devices use digital signal processors ("DSP") and other devices with limited computing power. In stark contrast iRecord uses a single Very Long Instruction Word ("VLIW") media processor running complex software completely developed in-house by Streaming Networks and refined over 8 years.

At the heart of iRecord is a robust real-time system including an H.264/AVC video encoder that delivers high coding efficiency with the highest video quality, and at the same time uses the computational resource optimally. There is no single coding element in Streaming Network's H.264/AVC encoder that provides the majority of the improvement in compression efficiency. Rather, it is the sum of numerous smaller improvements and innovations made by a team of 60 engineers over eight years that add up to the significant gain. The areas targeted by Streaming Networks H.264/AVC encoder with patent pending techniques are that of the rate controller, in-loop deblocking filter, efficient motion estimation, and the context-based arithmetic coding scheme.

The H.264/AVC video standard supports a base-line entropy coding method known as Context Adaptive Variable Length Coding ("CAVLC") and a high efficiency entropy coding method known as Context-based Adaptive Binary Arithmetic Coding ("CABAC"). CABAC typically provides a 10-15% bitrate reduction over to CAVLC when encoding TV signals at the same quality. However, CABAC's inherent sequential nature makes it computational demanding. For iRecord, the VLIW DSP-CPU load was controlled using patented technique of increased instruction level parallelism in the CABAC encoding process, reducing function call overhead, and increasing the efficiency of bit stream writing and other critical processes.

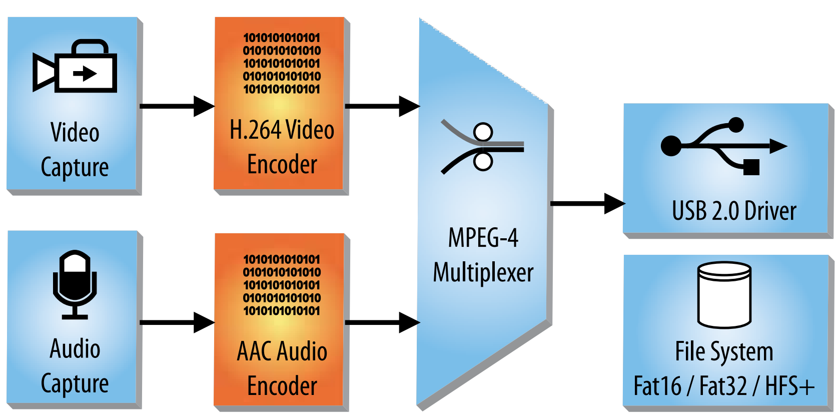
H.264/AVC video standard supports multiple block sizes ranging from 16x16 to 4x4 for accurate motion compensated prediction. As the block size becomes small the predicted residual data reduces but the amount of motion vector data to be compressed increases. The motion estimation process has to search for the best prediction mode to maximize compression and hence is highly demanding. Streaming Networks developed an extremely efficient motion estimator that distributes the overall burden of motion estimation among the DSP-CPU and other on-chip processing elements.

An unfortunate characteristic of block-based video coding is visible “blocking” artifacts created when block edges are typically reconstructed with less accuracy than inner pixels. To avoid this issue H.264/AVC defines an adaptive in-loop deblocking filter, which reduces the blockiness without affecting the sharpness of the content too much. This results in significant improvement in the video quality, while simultaneously reducing bit-rate up to 10% as well. Hence, it is a valuable tool to have in a H.264/AVC encoder. Streaming Networks has exploited fine grain parallelism in the in-loop deblocking filter algorithm to increase the instruction level parallelism of the VLIW DSP-CPU for computational gain.

The issue of video encoder rate and distortion control has gained significance with the arrival of H.264/AVC, since it is largely declarative which creates many more coding options, and hence opportunities, than previous standards. The H.264/AVC encoder has to select between numerous inter and intra-macroblock prediction modes to obtain the optimum encoding mode. This is a critical and time-consuming step, and the impressive compression of H.264/AVC largely depends on it. Streaming Networks has developed a patent-pending low cost high-efficiency encoder rate and distortion control mechanism, which utilizes a combination of feedback control techniques and an understanding of the factors that influence human visual perception.

Streaming Network’s numerous H.264/AVC optimizations, coupled with an intimate knowledge of the underlying hardware, enable iRecord to encode video at 30 fps up to D1 resolution in real-time in addition to audio encoding. The 2 audio channels are sampled at 44.1/48 KH$_z$ and encoded in AAC format at 128 Kbps. The audio and video is multiplexed in MP4 file format and written to the PMP or storage device using the USB 2.0 interface.

Figure 2: iRecord video recording application



Source: Streaming Networks

The transcoding application of iRecord Pro is computationally even more demanding then the recording only application. It decodes a D1 resolution (720 x 480) MPEG-2 files and encodes them into QVGA H.264 format in real-time.

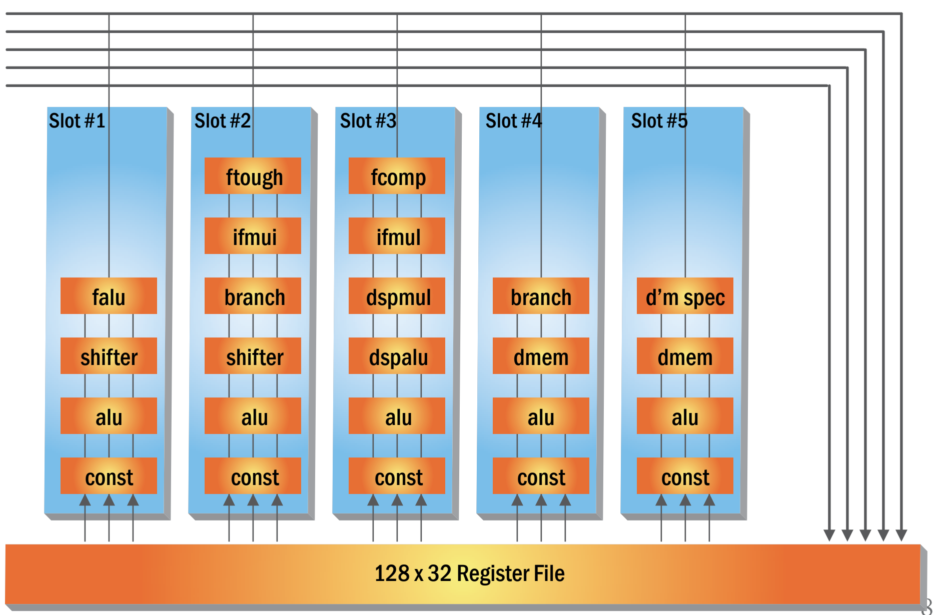
The processor used in iRecord is a 500 MHz, 5-slot VLIW media processor from NXP Semiconductors. The Nexperia™ processor includes on-chip video, audio and data I/O and multi-media operations units, beside the 32-bit VLIW DSP-CPU. These on-chip processing elements such as an MPEG-2 Variable Length Decoder (VLD) takes load off the DSP-CPU. Furthermore, the VLIW instruction set includes special custom operations to assist media processing intensive computations. It is the optimal use of these media processing custom operations that enables managing the very demanding motion compensation and CABAC modules of the H.264 video encoder in iRecord.

Even the most advanced DSPs do not have specialized instructions to perform video processing. In many cases, a number of instructions, or even sequences of DSP operations, are required to perform a given video function.

Custom Instructions Push Video Processing

The VLIW instruction length allows five simultaneous operations to be issued every clock cycle. These operations can target any five of the tens of functional units in the DSP-CPU core, including integer and floating-point arithmetic units and data-parallel multimedia operation units.

Figure 3: A simplified diagram of internal structure of a Nexperia VLIW core



Source: NXP and Streaming Networks

These multimedia operations significantly accelerate standard video and audio compression and decompression algorithms. As just one of the five operations issued in a single VLIW instruction, a single custom or media operation can implement up to 11 traditional microprocessor operations. These custom or multimedia operations combined with the VLIW architecture yield considerable throughput for video applications such as iRecord Pro.

Custom operations can substantially increase the processing speed in small kernels of applications. Their power comes from their ability to operate on multiple data items in parallel, basically Single Instruction, Multiple Data ("SIMD") operations. Hence, careful use of custom operations has the potential to reduce the absolute number of operations needed to perform a computation.

VLIW processing provided by Nexperia series of processors from NXP is vital for high resolution encoding a live event, which is highly demanding since the encoder does not have an opportunity to execute computational intensive compression algorithms in a different sequence. To achieve the best compression, the encoder must keep up with the incoming datastream. In such instances fast hardware, a VLIW processor, and optimized software are critical to delivering a robust real-time system.

streamingnetworks
digital video delivered