A NEW GENERATION OF MPEG-2 VIDEO ENCODER ASIC & ITS APPLICATION TO NEW TECHNOLOGY MARKETS

M Irvin, T Kitazawa, T Suzuki

Sony Broadcast & Professional Europe , UK

ABSTRACT

The multimedia era is upon us. The traditional methods used for the storage, signal processing and distribution of video, audio and data are going through significant change. This change is being fuelled by recent advances in the area of bit-rate reduction technology and in particular the publication of an effective world-wide standard for the compression of video and audio components, the MPEG-2 specification (1).

This paper considers the complexities involved in performing video compression encoding. It introduces newly developed Sony programmable ASIC's for performing real time compression encoding of a video source image in compliance with the MPEG-2 Main Profile @ Main Level standard. It details the architecture of this newly developed chipset including features that have been implemented within the chipset to identify image complexities, resulting in efficient video compression encoding with high picture quality.

The paper describes how this video compression chipset has been integrated into a new generation MPEG-2 Encoder product from Sony and identifies a number of new technology application areas which are being addressed by this new product.

INTRODUCTION

Video compression technology has made significant advances in recent years and a succession of compression formats have emerged. Figure 1 identifies a number of such compression formats and compares their achievable data rate range.

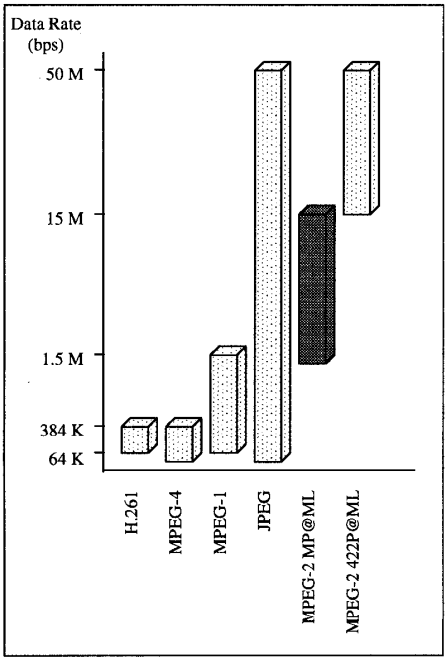

Figure 1 Compression Formats

Although each codec has its application area, one of the most prominent is the MPEG-2 Main Profile @Main Level because of its widely covered data bit-rate and sophisticated reduction method.

MPEG-2 VIDEO COMPRESSION TECHNOLOGY

The compression technique adopted by the Motion Picture Experts Group for the MPEG-2 family of profiles and levels is designed around an asymmetrical process. The nature of this asymmetrical process is designed to concentrate the required complex image analysis and processing to the compression encoding stage, whilst realising a significantly less complex decompression decoding stage.

By adopting this asymmetrical process, a cost effective decoding device can be achieved, although this does result in a cost penalty for the encoding device in comparison. However when considering a typical application area for MPEG-2, it can be seen that the number of decoding devices required far exceeds the number of encoding device by many thousand or hundred of thousands. The encoding device cost penalty can therefore be sustained.

The MPEG-2 specification is defined with respect to the decoding process and uses what is termed a 'System Target Decoder'. It is therefore open to manufacturers of MPEG-2 encoding systems to utilise to their best advantage techniques to improve the encoding process, as long as the resultant bit-stream is both syntactically correct in terms of MPEG-2 and that the resultant bit-stream can be decoded by the MPEG-2 defined System Target Decoder.

International Broadcasting Convention, 12–16 September 1996
Conference Publication No. 428, © IEE, 1996

We can summarise this by saying that in terms of picture quality, there is little perceived difference between different manufacturer supplied decoder devices. Moreover it is the encoding techniques adopted by a particular manufacturer which achieve the overall resultant picture quality and hence positions a manufacturers MPEG-2 encoding device against its competitors.

MPEG-2 VIDEO ENCODING CONSIDERATIONS

Within a video compression application, the source picture to be compressed can be found to have a number of characteristics. These characteristics are dependent upon:-

- The method used to capture or record the original source image, and
- The content of the original scene or image that was captured

By identifying these characteristics, a video compression encoding device can exploit the rich set of tools that the MPEG-2 syntax provides to produce are more efficient coded image.

Image Capture -The method used to capture the original source image will determine the nature of the picture type, for example, an image captured using a film type camera will determine a progressively scanned frame, whereas an image captured using a video type camera may determine an interlaced frame consisting of two fields.

In addition to the method used to capture the original image, we must also consider how the original image was recorded and 'transported' throughout the production and post production chain. In the case of an image that was captured using a film type camera, it is possible that the film image sequence may have been transferred to a video based format which has both a different frame rate and frame structure.

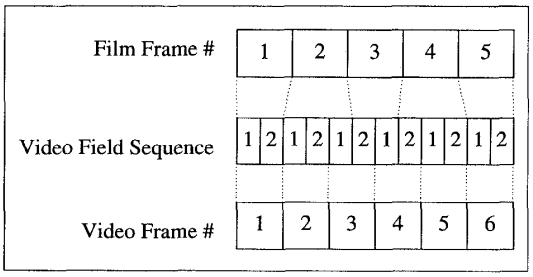

Figure 2 NTSC 3:2 Pulldown Sequence

In NTSC television based countries, it is common to transfer 24 frame per second film based material to a video based storage device, which characterises both an interlaced frame picture structure and a field rate of 59.94 Hz. The resultant transfer process is referred to as '3:2 pulldown'. This technique is detailed in Figure 2, in which it can be seen that the progressive frame structure is firstly converted to an interlaced field structure and secondly each fifth video field is a repeat field to achieve the desired overall video field rate.

In PAL based countries, the situation is slightly easier, as the conversion process involves a transfer from 24 frame film to an interlaced 50 field based video system. This is commonly achieved by increasing the replay rate of the film source material to 25 frames per second and then performing a progressive to interlaced scan conversion, resulting in two interlaced video fields which represent the original single film frame.

One cause for concern however, in PAL based film transfers, is the video field alignment accuracy. It is possible within the film transfer process, that the resultant interlaced video field to frame relationship is disturbed. This causes the odd interlaced video field and the even interlaced video field in a video frame to have originated from different consecutive film frames. This error process is commonly referred to as 'PAL field flip'.

Having identified the source picture image type, the MPEG-2 syntax provides tools which can be utilised in a manner that efficiently encodes the input video sequence. For example at the macroblock layer, the flags representing dct_type and prediction can be described as either field or frame based. In the case that an input video sequence has been 3:2 pulldown processed, a video encoding device which correctly identifies the 3:2 sequence can use the repeat_first_field flag at the picture coding layer and hence avoid the need to encode redundant fields, leading to increased encoding efficiency.

Image Content -The image content of the video sequence to be encoded can have a drastic effect on the overall picture quality. In everyday use, we can see video material which contains images representing movie film, sports events, commercials and news for example. Each image type has varying characteristics of image complexity placing differing demands on the video encoding process. Complexities can be dependent on scene content, such as fast moving sports images or strobe lighting, or on the image sequence structure itself, such as scene changes or fade-up and fade-down effects.

It is important that a video encoding device is able to dynamically react to these differing image complexities and keep encoding artifacts to a minimum. By proficient use of the MPEG-2 coding syntax and the resultant effective use of macroblock parameters, derived from inspecting each macroblock statistical characteristic, this can be achieved.

In the case of Constant Bit Rate (CBR) applications, additional constraints apply to the video encoding

process. The document TM-5 (2), Test Model 5 of the MPEG-2 standard gives guidance in this area. TM-5 determines three strategies to be adopted to balance the resultant video encoded bit-rate:-

- By determining each picture type I, P, or B complexity and the balance of bit allocation by feedback from the last picture of the same type.
- The use of a quantisation index, which is controlled each macroblock using feedback from the Variable Length Coding (VLC) buffer.
- By controlling the macroblock quantisation to improve picture quality dependent on macroblock activity.

It is worth noting however that TM-5 does not guarantee Video Buffer Verifier (VBV) compliance. Therefore to prevent underflow or overflow on VBV, zero stuffing and skip macroblock or skip picture syntax code is usually used.

SONY MPEG-2 VIDEO ENCODING SOLUTIONS

For MPEG-2 based applications, Sony has developed a two chip Application Specific Integrated Circuit (ASIC) for real time video encoding. The devices use 0.35 micron fabrication technology with a gate count approaching 1 million gates per device. This has achieved low power consumption and small size.

Table 1 details further the Sony device specifications.

	ENCONT	ENCORE
Technology	0.35 µm	0.35 µm
Gate count	910 K	910 K
Package	304 SQFP	208 SQFP
Power supply	2.5 V	2.5 V
Power consumption	0.8 W	0.8 W
External DRAM	4 ~ 32 MB	2 ~ 8 MB

Table 1 Specification of Chipset

The chipset consists of two devices designated the names ENCONT and ENCORE whose architecture is shown in Figure 3 . The ENCONT, encoder control device performs input signal adaptation and motion estimation, whereas ENCORE, the encoding core performs the compression process.

In designing the chipset, a hybrid architecture has been adopted. High speed and stereotype signal processing is assigned to hardware and complex algorithmical processes such as adaptive control and rate control are assigned to programmable RISC processors.

ENCONT -The ENCONT device performs frame re-ordering, motion estimation control, picture statistics and adaptive control. ENCONT can also perform the following operations:-

- Automatic detection of a 3:2 pulldown sequence and effective allocation of the repeat_first_field flag in the picture coding header.
- In the case of 625/50 video line standard operation, ENCONT can detect and compensate for PAL field flipped telecine film sources material.
- Automatic scene change detection, to improve picture quality by dynamically changing picture type and bit allocation.

ENCORE -The ENCORE device is responsible for Discrete Cosine Transform (DCT), Quantisation, Variable Length Coding (VLC) and local decoding. The on chip RISC processor within ENCORE is responsible for both constant and variable bit-rate control, including VBV compliance and has the capability to improve upon TM-5 rate control. The processor is programmable to insert all of the required MPEG-2 header syntax.

The chipset has selectable external memory size, such that for low_delay applications it can minimise memory size operation to 6 MB total.

The chipset can encode MPEG-2 Main Profile @ Main Level up to 15 Mbps, not only in real time with constant bit-rate, but also as a two pass encoding process with variable bit-rate as required for Digital Video Disk (DVD) authoring.

In other MPEG-2 application areas, such as Digital Video Broadcasting (DVB), both constant bit-rate and dynamically variable bit-rates, for Statistical Multiplexing operation can be supported.

By implementing a RISC processor inside each encoder chip, a variety of MPEG-2 applications can be realised by updating the executable microcode on the RISC core.

The Sony BDX-E1000 MPEG-2 Encoder is the first product to feature this newly developed Sony MPEG-2 encoding chipset ENCONT and ENCORE. The re-programmable nature of the encoding chipset at the heart of the BDX-E1000, gives provision to address a number of application areas with a single product development. This paper will now highlight some of the new application areas which are being addressed by the BDX-E1000.

APPLICATIONS OF MPEG-2

Digital Video Broadcasting

Digital Video Broadcasting (DVB) is the European standard for next generation programme transmission. The standard was developed with co-operation from over 150 members of a consortium representing a broad spectrum of the broadcast industry.

The DVB standard encompasses all elements of the broadcast chain, from studio playout to home reception. Source programme information, such as video, audio and teletext information is transmitted in a digital form.

To achieve efficient use of spectrum bandwidth, the DVB committee have endorsed the use of the MPEG-2 compression standard and in particular Main Profile @ Main Level video compression. The utilisation of compression enables a number of programme channels to be transmitted in the same spectrum space previously reserved for a single analogue channel.

The key elements of a Digital Video Broadcasting system are highlighted in Figure 4. Within the system, source programming is delivered from the programme playout area to the input of a number of MPEG-2 encoders. A single MPEG-2 encoder is designated to process the component parts of a single programme channel.

The resultant compressed bit-stream output from each of the MPEG-2 encoders is combined in to a single data stream by a Multiplexing device, the output of which is subsequently transmitted across a network. The DVB have currently defined four network types for the distribution of compressed programme, satellite, cable, terrestrial and microwave.

To complete the DVB transmission chain, a decoder unit is required at the consumers home to receive the transmitted signal. This device, referred to as an Integrated Receiver Decoder (IRD) performs signal adaptation, demodulation and MPEG-2 compliant Main Profile @ Main Level decoding.

The Sony BDX-E1000 MPEG-2 Encoding unit, which features the new generation video ASIC's described in this paper is fully DVB compliant. The unit has a flexible architecture to meet differing broadcasters requirements, making it an ideal product for DVB application.

The unit is compact in size and is designed to process all of the component parts of a single programme channel, including video, up to 8 audio channels and private data such as DVB subtitles. The device can be controlled and configured from an external PC or Workstation using time schedule information through either Ethernet or RS-232C connection.

Program Archive

The popularity of the television image within the last 30 years, has created an explosion in the television programme production market. This has resulted in the creation of vast libraries containing historical archives of programme material. As the popularity of television continues to grow, particularly as we are now moving into the age of digital television, there is a constant need for programme material to fill airtime schedules.

One of the potential sources for this programme material is from the program archive collections which contain many classic television programme produced within the most popular period of television history to date.

Unfortunately, many of these programme's were produced on storage media, which is now becoming unstable as it nears the end of its life. To prevent irretrievable loss of this precious programme material, the archivist must regenerate this material on to a new storage media. However this can be a very time consuming and costly process as new storage media has to be purchased.

The application of MPEG-2 compression to Program Archive can bring significant benefits, by offering a flexible compression scheme which can be optimised to meet picture quality against storage bandwidth requirements.

Figure 5 details a Program Archive system, in which a number of MPEG-2 Main Profile @ Main Level Encoders are located at the heart of the system. Programme source material is played into the MPEG-2 Encoders, which is subsequently compressed and stored on a new storage device, in this example a data tape recorder located in a robotic cart machine.

The archive process is controlled by a higher level software application, from which the archivist creates a database record of the archive material and determines the parameters to be used for the compression process.

The Sony BDX-E1000 MPEG-2 Encoder is ideally suited to the Program Archive application, as it supports numerous input video and output data interface formats, meeting customer requirements for choice of storage devices.

Digital Video Disk

The Digital Video Disk (DVD) is the next generation of consumer optical media, offering vastly increased storage capacity over its predecessors, with the ability to store an entire programme movie in digital video format on a single disk side.

To achieve this storage capacity, DVD has adopted MPEG-2 Main Profile @ Main Level compression. The DVD format gives provision for data rates averaging 3.5 Mbps, with peak demands as high as 10 Mbps. Higher bit-rates are assigned to complex pictures and lower bit-

rates to simpler pictures using an adaptive variable bit-rate process. This process is performed in two stages, the first being an evaluation stage to determine picture complexity and the second the actual compression process.

The newly developed ASIC's described in this paper can support this two stage coding process. Figure 6 details a Digital Video Disk Mastering system, in which the Sony BDX-E1000 MPEG-2 Encoder product forms an integral part.

The source programme material components are passed through their respective encoding devices and then combined at a multiplexing and disk formatting stage. The compressed and formatted disk master data is then streamed on to a removable media located inside a data storage device. This removable media is then forwarded to the disk mastering plant, where the optical disks are subsequently pressed.

CONCLUSION

This paper has described SONY's new generation MPEG-2 video ASIC's and their application to new technology markets.

High picture quality is achievable in both constant bit-rate and variable bit-rate compression modes through the use of image type recognition functions and adaptive bit-rate control by the video ASIC's.

The paper described three application areas addressed by a new Sony product featuring the new generation video ASIC's, Digital Video Broadcasting, Program Archive and Digital Video Disk.

REFERENCES

1. ITU-T Recommendation H.262 | ISO/IEC 13818:1995, Information technology -Generic coding of moving picture and associated audio information.

2. International Organisation for Standardisation, ISO/IEC JTC1/SC29/WG11 Coded Representation of Picture and Audio information. ISO/IEC JTC1/SC29/WG11/N0400 Test Model 5.

ACKNOWLEDGEMENTS

The authors would like to thank their colleagues for their contributions and support during the preparation of this paper.

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.