Single-chip MPEG-2 422P@HL CODEC LSI with Multi-chip Configuration for Large Scale Processing beyond HDTV Level

Hiroe Iwasaki Jiro Naganuma Koyo Nitta Ken Nakamura Takeshi Yoshitome
Mitsuo Ogura Yasuyuki Nakajima Yutaka Tashiro Takayuki Onishi Mitsuo Ikeda
Makoto Endo
NTT Cyber Space Laboratories, NTT Corporation
1-1 Hikarinooka, Yokosuka-Shi, Kanagawa Pref. 239-0847, Japan
iwasaki.hiroe@lab.ntt.co.jp

Abstract

This paper proposes a new architecture for VASA, a single-chip MPEG-2 422P@HL CODEC LSI with multi-chip configuration for large scale processing beyond the HDTV level, and demonstrates its flexibility and usefulness. This architecture consists of triple encoding cores, a decoding core, a multiplexer/de-multiplexer core, and several dedicated application-specific hardware modules with a hierarchical flexible communication scheme for high-performance data transfer. VASA is the world's first single-chip full-specs MPEG-2 422P@HL CODEC LSI with a multi-chip configuration. The VASA implements MPEG-2 video and system CODEC with generic audio CODEC interfaces. An LSI incorporating the architecture was successfully fabricated using the 0.13-μm eight-metal CMOS process. The architecture not only provides an MPEG-2 422P@HL CODEC but also large scale processing beyond the HDTV level for digital cinema and multi-view/-angled live TV applications with a multi-chip configuration. The VASA implementations will lead to a new dimension in future high-quality, high-resolution digital multimedia entertainment.

1. Introduction

Recent progress in video and audio compression technology has made it possible to provide a much greater volume and range of digital multimedia. The MPEG-2 standard [1] has emerged as a method for effectively compressing video and audio with high quality. In particular, this standard is currently seeing extensive worldwide use in many transmission and storage applications, such as digital satellite broadcasting, digital terrestrial broadcasting, digital cable television, video conferencing, DVD and CD-ROM storage media, video on demand, and time-shifted viewing.

The digitization of TV broadcasting is descending upon the world in the form of a global wave. This can be seen, for example, in the advent of terrestrial digital broadcasting, which will be offered in Japan by the end of 2003. With this broadcasting technique, producing programs that is the form of MPEG-2 transport stream and exchanging them over broadband digital networks will also boost their circulation. For this scenario to become a reality, it will be necessary to develop HDTV CODEC systems that are not only of professional quality but also compact. Towards this end, we have developed a number of HDTV encoder systems, the most recent and smallest of which is an 1U-half-rack, 9-chip HDTV system [2]. However, small-board, single-chip systems will be required for the future.

In recent years, several single-chip MPEG-2 encoder LSIs [4, 5, 6, 7, 8, 3, 9] have been developed to implement MPEG-2 codecs, but generally they are only able to perform MP@ML or 422P@ML video encoding. Some of them [3, 8, 9] also provide 422P@HL video encoding with a multi-chip configuration. Several single-chip MPEG-2 decoder LSIs for HDTV have also been developed [10, 11]. We have also developed MPEG-2 video encoder LSIs [12, 3] and have implemented compact and high-quality MPEG-2 CODEC systems [13, 14, 15, 2].

There are two major problems/requirements involved in implementing single-chip HDTV CODEC LSI, as follows:

- *Complexity and memory bandwidth:* The MPEG-2 422P@HL video encoding in a HDTV CODEC system requires a large amount of computational complexity for processing and large memory bandwidth for charging data as a factor of six times compared to those of 422P@ML encoding for SDTV respectively. Parallel encoding with its each memory has been well studied [3, 8, 9], but no architecture for a single-chip HDTV CODEC exploiting parallel encoding cores and a very high-speed external memory to solve these requirements has yet been developed.
- *Multi-chip configuration:* This will be necessary to provide futuristic high-quality, high-resolution digital multimedia entertainment beyond the HDTV level, such as digital cinema and multi-view/-angled live TV for broadcasting sporting events, music events and the like. Capabilities of a multi-chip configuration for

large scale processing beyond the HDTV level can not be disregarded.

To solve these problems, we propose a new architecture for a single-chip MPEG-2 422P@HL CODEC LSI with multi-chip configuration for large scale processing beyond the HDTV level, and demonstrates its flexibility and usefulness. This architecture consists of triple encoding cores, a decoding core, a multiplexer/de-multiplexer core, and several dedicated application-specific hardware modules with a hierarchical flexible communication scheme for high-performance data transfer. This chip, named VASA, implements MPEG-2 video and system CODEC with generic audio CODEC interfaces. An LSI incorporating the architecture was successfully fabricated using the 0.13-μm eight-metal CMOS process. The architecture not only provides an MPEG-2 422P@HL CODEC but also large scale processing beyond the HDTV level for digital cinema and multi-view/-angled live TV applications with a multi-chip configuration.

Structure of the paper. Section 2 describes the LSI architecture, and Section 3 explains how it was implemented in the fabricated LSI and evaluates the resultant LSI's performance. Section 4 presents the multi-chip configuration and its applications.

2. Architecture

2.1. Approach

The VASA architecture is derived from a re-modeling of "Parallel Encoding", a modification from our previous parallel encoding model for the portable HDTV encoder [2], i.e., from "an individual address space model" to "an unique address space model".

The control and data hierarchy of parallel encoding is based on macroblock pipelined schemes in each parallel encoding core (intra-core:second level) and inter-core (intra-chip:top level). And, this hierarchy is also the same as the two-level memory hierarchy for intra- and inter-core. The control and data hierarchy is mapped on that of the VASA architecture, named HFCA (hierarchical flexible communication architecture).

HFCA is like dual hierarchical backbones linked to every module and core in a chip and every sub-module in a core. One backbone is for small amounts of control data linked to the CPU-BUS, and the other is for huge amounts of picture data linked to the System-BUS.

2.2. Hardware Architecture

2.2.1 Block diagram

The block diagram of VASA is shown in Figure 1. The VASA consists of a RISC processor(TRISC), triple video encoding cores(E-CORE) each with a RISC processor(VRISC), a hardwired decoding core(D-CORE), a multiplexor/de-multiplexor core(MDX) with a RISC processor(SRISC), a video interface/display(VIF/DISP), a

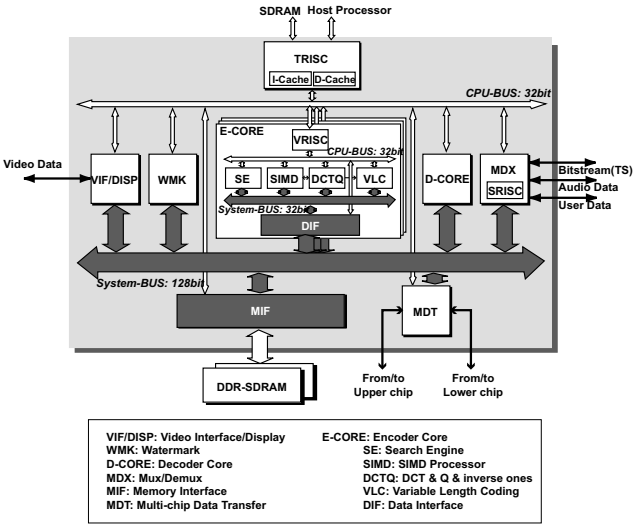

Figure 1. Block diagram of VASA.

water-mark(WMK), a multi-chip data transfer(MDT), and a memory interface(MIF).

All application-specific hardware modules and cores dedicated to the MPEG-2 implementation are connected to both a hierarchical "CPU-BUS" and a hierarchical "System-BUS." The backbone for the control hierarchy is controlled by the TRISC and the three VRISCs in each "E-CORE", with a hierarchical "CPU-BUS" structure. The other backbone for the data hierarchy is controlled by the MIF and a data interface module (DIF) in each "E-CORE," with a hierarchical "System-BUS" structure.

2.2.2 Hierarchial flexible communication scheme

The HFCA provides space and time switching between inter- and intra-chip communication and inter- and intra-core communication paths on which data can be transferred either immediately, or after a certain time interval between any of the chips, cores, modules, and sub-modules in the same manner.

The MIF and the DIFs, which are a hierarchical "System-BUS" structure, are programmable controllers with instructions, and the operation of the MIF and the DIFs can be changed by modifying the instructions. The MIF controls the external DDR-SDRAMs in order to minimize their I/O rate. The HFCA makes it possible to increase the average ratio of the active bandwidth of the DDR-SDRAMs to 70% in ordinal encoding and to 85% in advanced encoding with pre-processing, thus ensuring high quality.

VASA's architecture provides sufficient performance and flexibility to serve the most demanding applications based on recent high-quality CODEC technologies.

A. Intra-core/intra-chip communication

The intra-core/intra-chip communication is shown in Figure 2. The flexible communication in a chip is performed by

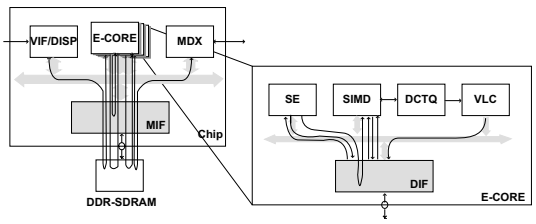

Figure 2. Intra-core/intra-chip communication.

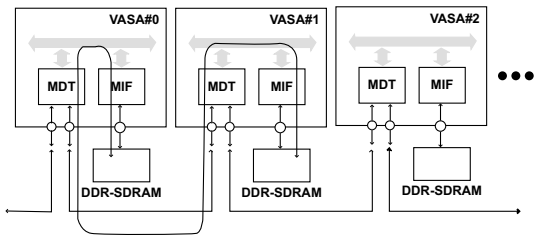

Figure 3. Inter-chip communication.

the TRISC and the MIF, while the flexible communication in a core is performed by the VRISCs and the DIFs. The picture data in a chip can be transferred immediately or after a certain time interval by the TRISC and the MIF; likewise, that in a core can be transferred immediately or after a certain time interval by the VRISCs and the DIFs. Therefore, spatial/temporal flexibility is achieved in the picture data transfer via the DIF and MIF software.

B. Inter-chip communication

The inter-chip communication is shown in Figure 3. The inter-chip data transfer through the MDT enables the VASA chips to use the data on other VASA chips. This scheme enhances multi-chip configuration scalability and thus enables large-scale processing beyond the HDTV level. The reference data between the VASA chips is composed of the sub-sampled original pixels and local decoded ones in the sub-pictures. The sub-picture data are transferred through each MDT in VASA, and are stored in each DDR-SDRAM.

2.2.3 Hierarchy macroblock pipeline control

The hierarchy macroblock pipeline control is shown in Figure 4. Data between DDR-SDRAMs and E-COREs are controlled by the MIF, and data in the E-COREs are controlled by the DIFs. The DIFs have double buffers for data-transferring, and the double buffers are switched in macroblock cycle. First, data from the DDR-SDRAMs to the application-specific hardware blocks (SE/SIMD, etc.) are transferred from the DDR-SDRAMs to the buffers of the DIFs. Second, the data are transferred from the DIF buffers to application-specific hardware blocks(SE/SIMD, etc.). The data from application-specific hardware blocks(SE/SIMD, etc.) are then transferred to the DDR-SDRAMs via the DIF buffers in the same manner. This hierarchical transferring is controlled by the MIF and DIFs in a macroblock cycle. This hierarchical macroblock pipeline control results in high data-transferring performance with sufficient flexibility.

2.3. Software Architecture

The software architecture is shown in Figure 5. It consists of three layers: the bottom is the hardware layer, which is the VASA hardware; the middle is the hardware control layer, which is TRISC and VRISC software for controlling the VASA hardware; the top is the function layer, which is TRISC software for handling MPEG-2 basic and user functions. Communication between the hardware layer and the hardware control layer is accomplished via the hardware/software interface in VASA. Communication between the hardware control layer and the function layer is accomplished via the function interface in TRISC.

The middle hardware control layer consists of TRISC's chip layer control and VRISC's E-CORE layer controller. Communication between TRISC and VRISC is achieved via the TRISC-VRISC interface. The top function layer consists of TRISC's common basic functions and the custom functions for users. Communication between common basic functions and the user custom functions is carried out via the custom function interface.

This software hierarchy not only completely frees users from having to tediously control VASA hardware using the low-level hardware/software interface and the difficulty of handling MPEG-2 common basic functions, but also provides a simple programming interface as a custom function. Thus, they can easily customize and concentrate on higher level programming to achieve high quality, high compression, and low delay based on the knowledge of CODECs.

3. Implementation and evaluation

3.1. LSI Characteristics

The physical features of the fabricated LSI are shown in Table 1. The VASA integrates 61.4 million transistors in a 14.0 mm x 14.0 mm chip using 0.13-μm 8-level metal CMOS technology. The clock frequency is 200-MHz. Power consumption is estimated to be 5.0-W at 200-MHz with 1.5-V to the internal circuits, 2.5-V to the DDR-SDRAM circuits, and 3.3-V to the I/O circuit supply source. The chip is mounted on a 1008-pin FCBGA. External memories are 256-Mbit (32-bit width) 200-MHz DDR-SDRAM x 2 for images and a 32-Mbit (16-bit width) 100-MHz SDRAM x 1 for the TRISC large software if necessary.

A micro-photograph of the VASA LSI is shown in Figure 6. As shown in the photograph, VASA integrates TRISC, triple E-CORES, D-CORE, MDX, VIF/DISP, WMK, MDT, and MIF. All logic circuits on the chip were

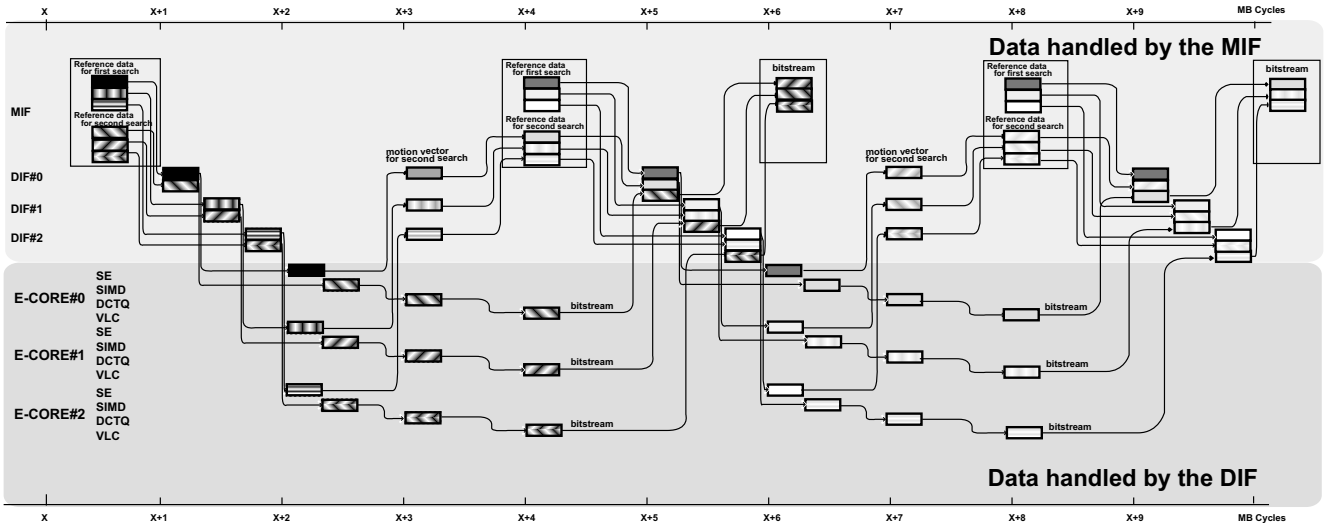

Figure 4. Hierarchy macroblock pipeline.

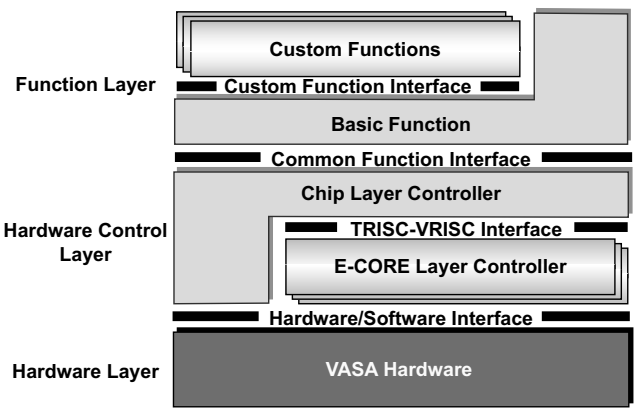

Figure 5. VASA software architecture.

Table 1. VASA physical features

Technology	0.13-μm 8-level metal CMOS
Number of transistors	61.4 million
Die size	14.0 mm x 14.0 mm
Clock frequency	200 MHz
Supply voltage	1.5 V/2.5 V/3.3 V
Power consumption	5.0-W(at 1080I 422P@HL)
Package	1008-pin FCBGA(35mm x 35mm)
External memories	256-Mbit(32-bit) 200-MHz DDR-SDRAM x 2(for images) and 32-Mbit(16-bit) 100-MHz SDRAM(for TRISC large software, *if necessary*)

constructed using only standard cells in order to shorten design time. Hard macros were not used except for the memories and the TRISC, which is a RISC core with I-/D- cache memories. Encoding and decoding cores on the chip used modified versions of our previous design cores.

The function features of the LSI are shown in Table 2. The MPEG-2 standards for the video encoding algorithms are supported by various profiles and levels with a single VASA chip. The resolution and frame rate supports 1920/1440 x 1080 at up to 30 frames per second with a single VASA chip. The maximum size and rate with VASA's multi-chip configuration are 4k x 2k at up to 60 frames/sec respectively. Moreover, the multi-view profile of MPEG-2 is also supported for the stereo image CODEC, and preprocessing for high-quality coding and water-marking for digital content protection are carried out in real-time. The system part I/O format and bitrate are supported by MPEG-2 TS (188/204 byte) and the maximum rate is as high as 300-Mbits per second. TS-multiplexing and demultiplexing are also supported for multi-chip configuration.

3.2. Software Characteristics

The software is written in C language except for the interrupt handler. The program size on TRISC is 35 Klines, and the sizes of both of the program and data memory for TRISC are 16-K words. If necessary, it uses external SDRAM as a chase. The TRISC controls VIF/DISP, WMK, MIF, MDX and E-CORE through VRISC in video encoding. The program size on VRISC is 11 Klines, and the sizes of both of program and data memory for VRISC are 6-K words. The software on the three VRISCs is the same, and VRISCs control SE, SIMD, DCTQ, VLC, and DIF in E-CORE. If necessary, it can use an external DDR-SDRAM as dynamic program replacing areas.

Table 2. VASA functional features

Video	
Profile and level	MPEG-2 422P, MP@HL, 422P, MP@H-14, 422P, MP, SP@ML
Search range	Narrow: -225.5/+211.5(H), -113.5/+125.5(V)
	Wide: -449.5/+435.5(H), -128.0/-127.5(V)
Resolution and rate	Single-chip: 1920/1440 x 1080 at up to 30 fps
	1280 x 720 at up to 60 fps
	Multi-chip: 4096 x 2048 at up to 60 fps(MAX)
Pre-processing	Macroblock based sophisticated function filter
Multi-view profile	Stereo image CODEC
Water-mark	Original water-mark insertion/extraction
Audio	
I/O format	Linear PCM or encoded stream(AAC)
User	
I/O format	PES format for time-code and other audio and data
System	
I/O format	MPEG-2 TS(188/204 bytes)
Bitrate	300 Mbps(MAX)
Multi-channel CODEC	Encoding/decoding by TS multiplexing/de-multiplexing

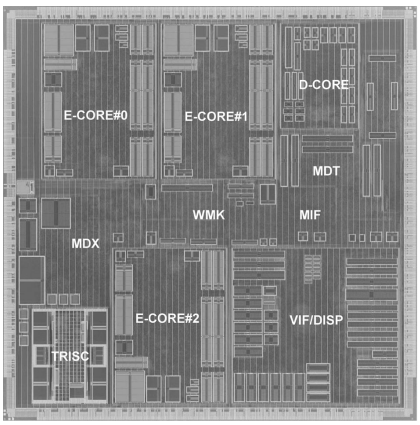

Figure 6. A micro-photograph of VASA LSI.

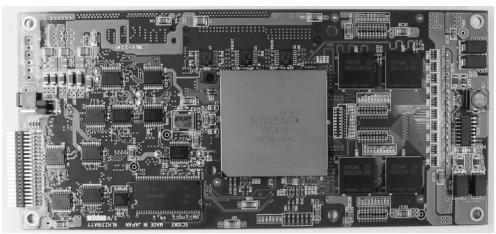

Figure 7. A VASA evaluation module.

3.3. Evaluation board

The evaluation module, which consists of the VASA, four 128-Mbit DDR-SDRAMs[1] with 32-bit bandwidth, video input, and TS output, is shown in Figure 7. The clock frequency on the board is 200 MHz. The evaluation module is very small, i.e., 15.8 cm x 7.2 cm. This evaluation module is combined with a CODEC mother evaluation board, and the various VASA functional features in Table 2 are tested on the evaluation board.

4. Multi-chip configuration and its applications

4.1. Multi-chip configuration

The VASA's multi-chip configuration provides larger size processing beyond HDTV for digital cinema and multi-angled live TV, multi-view profile for stereo image CODECs, and multi-channel CODECs with TS multiplexing and de-multiplexing.

A typical multi-chip system configuration is shown in Figure 8. It has some 422P@HL encoders, each of which consist of VASA and DDR-SDRAMs. The number of sets of 422P@HL encoders is dependent on the required resolution and frame rate. The neighbor 422P@HL encoders are connected to each other in a two-way manner. One is inter-chip communication using MDT, and the other is daisy-chained using the TS-multiplexer function. The bottom 422P@HL encoder outputs a single multiplexed MPEG-2 transport stream (TS) without any extra circuits or equipment.

4.2. Multi-chip applications

The first multi-chip application is a digital cinema. Original super high-definition images beyond the HDTV level composed of images such as 4k x 2k pixels are horizontally divided into several pieces of proper sub-pictures (same as HDTV level), and these sub-pictures are allocated to several 422P@HL encoders including the VASA chip. After parallel encoding on the VASA multi-chip configuration, a TS of digital cinema is output from the bottom of the VASA chip.

The second multi-chip application is multi-view/-angled live TV for sporting events. Original multi-angled TV images beyond the HDTV level, which consists of several

[1]The requirement for DDR-SDRAM in this LSI is 512-Mbit capacity and 64-bit bandwidth.

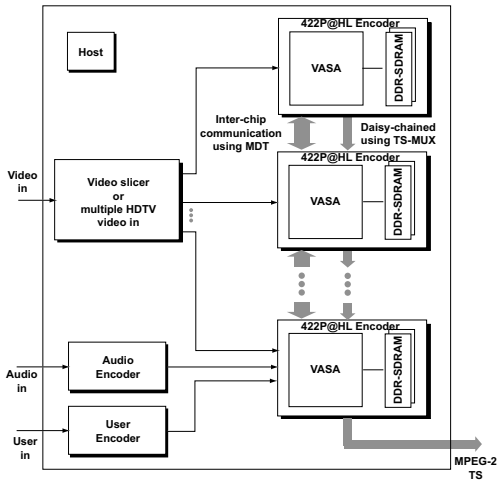

Figure 8. A multi-chip system configuration.

HDTV images such as a batter, catcher, and third baseman etc., are allocated to several 422P@HL encoders including the VASA chip. After parallel encoding on the VASA multi-chip configuration, a TS of multi-view/-angled live TV is output from the bottom of the VASA chip.

In these way, VASA's multi-chip system configuration can handle large scale processing beyond the HDTV level easily in real-time. These futuristic high-quality and high-resolution digital multimedia entertainment will become an important and popular feature in people's lives in the near future.

5. Conclusion

This paper proposed a new architecture for a single-chip MPEG-2 422P@HL CODEC LSI with multi-chip configuration for large scale processing beyond the HDTV level, and demonstrates its flexibility and usefulness. This architecture consists of triple encoding cores, a decoding core, a multiplexer/de-multiplexer core, and several dedicated application-specific hardware modules with a hierarchical flexible communication scheme for high-performance data transfer. The VASA implements MPEG-2 video and system CODEC with generic audio CODEC interfaces. An LSI incorporating the architecture was successfully fabricated using the 0.13-μm eight-metal CMOS process. The architecture not only provides an MPEG-2 422P@HL CODEC but also large scale processing beyond the HDTV level for digital cinema and multi-view/-angled live TV applications with a multi-chip configuration.

References

[1] ISO/IEC 13818-1/2/3 International Standard, *Information Technology – Generic Coding of Moving Pictures and Associated Audio: Systems/Vidual/Audio* –, November 1994.

[2] T. Yoshitome, K. Nakamura, K. Nitta, M. Ikeda, and M. Endo, "Development of an HDTV MPEG–2 encoder based on multiple enhanced SDTV encoding LSIs," in *International Conference on Consumer Electronics(ICCE) 2001*, 2001, pp. 160–161.

[3] M. Ikeda, T. Kondo, K. Nitta, K. Suguri, T. Yoshitome, T. Minami, J. Naganuma, and T. Ogura, "An MPEG-2 Video Encoder LSI with Scalability for HDTV based on Three-layer Cooperative Architecture," in *Design, Automation and Test in Europe Conference 1999*, March 1999, pp. 44–50. Also available in T. Minami, T. Kondo, K. Nitta, K. Suguri, M. Ikeda, T. Yoshitome, H. Watanabe, H. Iwasaki, K. Ochiai, J. Naganuma, M. Endo, E. Yamagishi, T. Takahashi, K. Tadaishi, Y. Tashiro, N. Kobayashi, T. Okubo, T. Ogura, and R. Kasai, "A single-chip MPEG2 MP@ML video encoder with multi-chip configuration for a single-board MP@HL encoder," in *HOT Chips X*. IEEE, August 1998, pp. 123–131.

[4] M. Mizuno, Y. Ooi, N. Hayashi, J. Goto, M. Hozumi, K. Furuta, A. Shibayama, Y. Nakazawa, O. Ohnishi, S. Y. Zhu, Y. Yokoyama, Y. Katayama, H. Takano, N. Miki, Y. Senda, and M. Yamashina, "A 1.5-W single-chip MPEG-2 MP@ML video encoder with low power motion estimation and clocking," *IEEE Solid-State Circuits*, vol. 32, pp. 1807–1816, November 1997.

[5] E. Miyagoshi, T. Araki, T. Sayama, A. Ohtani, T. Minemaru, K. Okamoto, H. Kodama, T. Morishige, A. Watabe, K. Aoki, T. Mitsumori, H. Imanishi, T. Jinbo, Y. Tanaka, M. Taniyama, T. Shingou, T. Fukumoto, H. Morimoto, and K. Aono, "A $100mm^2$ 0.95w single-chip MPEG2 MP@ML video encoder with a 128GOPS motion estimator and a multi-tasking RISC-type controller," in *International Solid-State Circuits Conference*, 1998, pp. 30–31.

[6] E. Ogura, M. Takashima, D. Hiranaka, T. Ishikawa, Y. Yanagita, S. Suzuki, T. Fukuda, and T. Ishii, "A 1.2w single-chip MPEG2 MP@ML video encoder LSI including wide search range motion estimation and 81MOPS controller," in *International Solid-State Circuits Conference*, 1998, pp. 32–33.

[7] S. Ishiwata, T. Yamakage, Y. Tsuboi, T. Shimazawa, T. Kitazawa, S. Michinaka, K. Yahagi, H. Takeda, A. Oue, T. Kodama, N. Matsumoto, T. Kamei, T. Miyamori, G. Ootomo, and M. Matsui, "A Single-Chip MPEG-2 Codec Based on Customizable Media Microprocessor," in *IEEE 2002 Custom Integrated Circuits Conference*, 2002, pp. 163–166.

[8] S. Kumaki, H. Takata, Y. Ajioka, T. Ooishi, K. Ishihara, A. Hanami, T. Tsuji, T. Watanabe, C. Morishima, T. Yoshizawa, H. Sato, S. Hattori, A. Koshio, K. Tsukamoto, and T. Matsumura, "A 99-mm^2 0.7-W Single-Chip MPEG-2 422P@ML Video, Audio, and System Encoder With a 64-Mb Embedded DRAM for Portable 422P@HL Encoder System," in *IEEE Journal of Solid-State Circuits*, 2002, pp. 450–454.

[9] LSI Logic Corporation., "(2002) LSI LOGIC home page.," [Online]. Available: http://www.lsilogic.com.

[10] H. Mizosoe, K. Maeda, Y. Kubo, Y. Tsuru, and K. Koruki, "An Advanced Multimedia Processing LSI Suitable for HDTV Applications," in *International Conference on Consumer Electronics(ICCE) 2001*, 2001, pp. 96–97.

[11] H. Yamauchi, S. Okada, K. Taketa, Y. Matsuda, T. Mori, S. Okada, T. Watanabe, Y. Harada, M. Matsudaira, and Y. Matsushita, "A 0.8W HDTV Video Processor with Simultaneous Decoding of Two MPEG2 MP@ML Streams and Capable of 30frames/s ReversePlayback," in *International Solid-State Circuits Conference*, 2002, pp. 473–474.

[12] T. Kondo, K. Suguri, M. Ikeda, T. Abe, H. Matsuda, T. Okubo, K. Ogura, Y. Tashiro, N. Ono, T. Minami, R. Kusaba, T. Ikenaga, N. Shibata, R. Kasai, K. Otsu, F. Nakagawa, and Y. Sato, "Two-chip MPEG-2 video encoder," *IEEE Micro*, vol. 16, pp. 51–58, April 1996.

[13] N. Terada, H. Ishii, T. Tachi, Y. Okumura, and H. Kotera, "A MPEG2-based digital CATV and VOD system using ATM-PON architecture," in *IEEE MULTIMEDIA 96 Conference*, 1996, pp. 522–531.

[14] Y. Tashiro, T. Izuoka, K. Yanaka, Y. Ito, N. Ono, Y. Yashima, H. Yamauchi, and H. Kotera, "MPEG2 video and audio CODEC board set for a personal computer," in *IEEE Global Telecommunications Conference*, 1995, pp. 483–487.

[15] M. Endo, J. Naganuma, Y. Nagajima, and T. Ogura, "An MPEG-2 Encoding PC Card System for Real-time Mobile Applications," in *International Conference on Consumer Electronics(ICCE) 2001*, 2000, pp. 242–243.