放送基盤向け HDTV 対応 H.264/AVC High422 プロファイル／MPEG-2 422 プロファイル符号化 LSI

新田　高庸[†]　池田　充郎[†*]　岩崎　裕江[†]　大西　隆之[†]　佐野　卓[†]
嵯峨田　淳[†]　中島　靖之[†]　稲森　稔[†*]　吉留　健[†]　松田　宏朗[†*]
谷田　隆一[†]　清水　淳[†]　中村　健[†]　長沼　次郎[†]

† 日本電信電話株式会社 NTT サイバースペース研究所
〒 239-0847 神奈川県横須賀市光の丘 1-1
＊ NTT エレクトロニクス株式会社
E-mail: †nitta.koyo@lab.ntt.co.jp

あらまし　放送基盤向け HDTV 対応 H.264/AVC High422 プロファイル符号化 LSI, "SARA/E" を開発した. SARA/E は, 現在, 放送素材伝送で用いられている MPEG-2 422 プロファイルにも対応している. 独自の 3 つの動き検出／動き補償エンジンにより, −217.75 から +199.75 (水平方向) / −109.75 から +145.75 (垂直方向) という広範囲の探索範囲を実現するとともに, H.264/AVC で規定されているほとんどすべての ME/MC 符号化ツールを利用可能である. 実験によると, 動きの速いシーンにおいて 1.2 dB から 1.7 dB の画質向上を達成している. SARA/E は 90 nm プロセスで 1 億 4 千万トランジスタを集積している.

キーワード　H.264/AVC, 符号化, MPSoC, 動き検出／動き補償, HDTV

An H.264/AVC High422 Profile and MPEG-2 422 Profile Encoder LSI for HDTV Broadcasting Infrastructures

Koyo NITTA[†], Mitsuo IKEDA[†*], Hiroe IWASAKI[†], Takayuki ONISHI[†], Takashi SANO[†],
Atsushi SAGATA[†], Yasuyuki NAKAJIMA[†], Minoru INAMORI[†*], Takeshi YOSHITOME[†],
Hiroaki MATSUDA[†*], Ryuuichi TANIDA[†], Atsushi SHIMIZU[†], Ken NAKAMURA[†],
and Jiro NAGANUMA[†]

† NTT Cyber Space Laboratories, Nippon Telegraph and Telephone Corporation
1–1, Hikarinooka, Yokosuka-shi, Kanagawa, 239–0847, Japan
∗ NTT Electronics Corporation
E-mail: †nitta.koyo@lab.ntt.co.jp

Abstract　An H.264/AVC encoder LSI (named "SARA/E") that supports High422 profile, as well as 422 profile of MPEG-2, has been developed for HDTV broadcasting infrastructures. It contains three motion estimation and compensation (ME/MC) engines with search ranges of −217.75 to +199.75 (H) / −109.75 to +145.75 (V), which can utilize almost all H.264/AVC ME/MC tools, multiple reference frame, variable block size, 0.25-pel prediction, macroblock adaptive field/frame prediction (MBAFF), temporal/spatial direct mode, and weighted prediction. Our evaluations show that it can encode fast moving scenes with 1.2 dB to 1.7 dB higher than the JM. It was successfully fabricated in a 90 nm technology. It integrates 140 million transisters.

Key words　H.264/AVC, encoder, MPSoC, ME/MC, and HDTV.

1. Introduction

The H.264/AVC [1] will play an important role in the field of HDTV broadcasting infrastructures, like DVB-H in Europe, ISDB-T in Japan, and US-ATSC. There are many professional applications, such as iterruption, contribution, and distribution. Requirements for encoder LSIs used in these high-end systems are as follows. 1) No "weak" scenes are permitted. Professional encoder LSIs should be able to encode a variety of scenes efficiently. 2) A 4:2:2 chroma format support for material contribution, which yields to increase 33 % memory bandwidth, is indispensable. 3) Additional functionality, such as transcoding or tandem (2-passed) encoding, is desired.

Although several consumer LSIs have already been developed [2]~[4], it is hard to implement a professional HDTV H.264/AVC encoder into a single chip even with a 90 nm technology. We, therefore, have developed a professional H.264/AVC video encoder LSI, SARA/E [5], that can be configured with multi-chip for HDTV. It is the successor to our previous MPEG-2 422P@HL CODEC chip (VASA) [6].

The SARA/E has a wide motion estimation and compenstation (ME/MC) with high precision in order to encode a variety of scenes efficiently. It also has an advanced coding control scheme (ACC), which is very useful to realize additional functionality. Memory mapping to reduce memory bandwitdh is also proposed.

This paper is organized as follows. Section 2. describes the system architecture of the SARA/E. The ACC and memory mapping are also argued in the section. The ME/MC architecture is explained in section 3. Two types of parallelism are introduced in an ME engine in order to expand the search area. Two kinds of SIMD processors are adopted for another ME engine. Implementation results are mentioned in section 4. Some image quality evaluations are shown in section 5. Then, section 6. is conclusions of this paper.

2. System architecture

2.1 SARA/E architecture

Fig. 1 shows a block diagram of the SARA/E. The MPSoC chip consists of a 64-bit RISC processor (TRISC), two video coding cores (M-CORE and C-CORE), a video interface (VIF), pre-analysis engines (IR, MBP, and RIT), a multiplexer (MUX) that can concatenate bitstreams of itself and ones from other SARA/E chips, a multi-chip data transfer (MDT) that can send/receive image data from/to other SARA/E chips, a memory interface (MIF), and embedded DRAMs (eDRAMs).

Each of the M-CORE and the C-CORE has a 32-bit RISC processor (MRISC and CRISC, respectively). The M-CORE

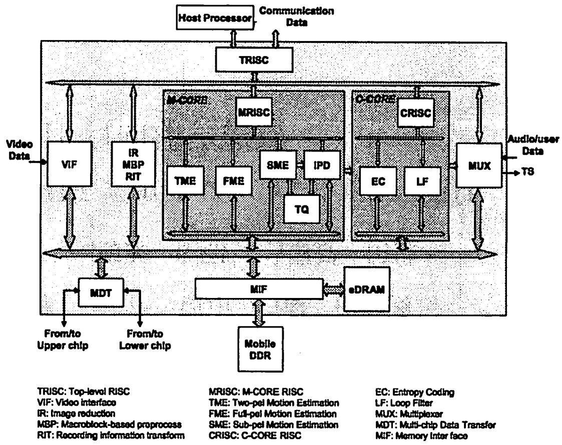

図 1 SARA/E アーキテクチャ.

Fig. 1 Block diagram of the SARA/E.

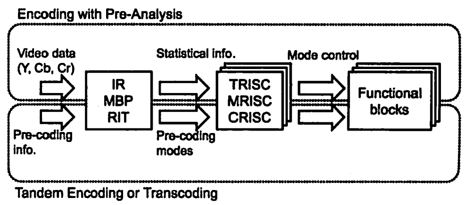

図 2 前処理を用いた符号化制御.

Fig. 2 Advanced coding control.

has triple ME/MC engines (TME, FME, and SME), an intra prediction (IPD), and a transform and quantization (TQ) as application-specific hardware modules. An entropy coding (EC) and a loop filter (LF) are in the C-CORE.

2.2 Advance coding control

The SARA/E has powerful pre-analysis engines (IR, MBP, and RIT). They can calculate statistical data of video signal, detect scene changes or fade scenes, and extract pre-coding modes. Together with the three RISC processors (TRISC, MRISC, and CRISC), they can realize an "advance coding control" (ACC) scheme (Fig. 2). Before encoding process, video data and pre-coding information, if any, are input into pre-analysis engines, which output statistical information and pre-coding modes to the three RISC processors. The RISCs can control functional blocks with various coding modes.

Using the ACC scheme, the SARA/E can encode fade scenes with automatic weighted prediction. Transcoding with inheriting pre-coding modes are also realized with the ACC.

2.3 Memory mapping

A 4:2:2 chroma format support increases the memory bandwidth because chroma data, Cr and Cb, are doubled. Although an eDRAM can reduce memory bandwitdh itself,

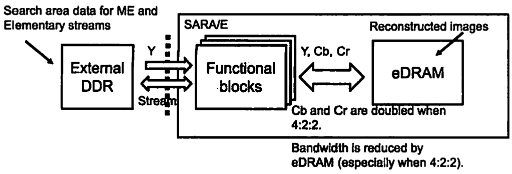

図 3　外部メモリと eDRAM 間のメモリマッピング.
Fig. 3　Memory mapping.

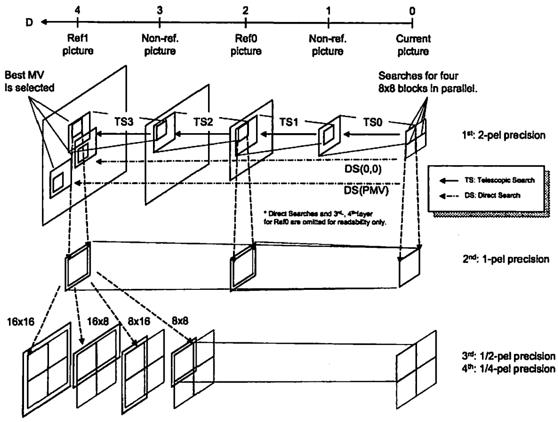

図 4　H.264 に拡張した動き検出/動き補償アルゴリズム.
Fig. 4　ME/MC algorithm for H.264/AVC.

memory mapping can reduce it more (Fig. 3). Search area data for ME are mapped into an external DDR-SDRAM, because no Cb and Cr data are needed. Reconstructed images are mapped into the eDRAM. This mapping can reduce bandwidth especially when 4:2:2 encoding.

3. ME/MC architecture

3.1 ME/MC algorithm

Our ME/MC algorithm is based on the one of [7], and extended and optimized for H.264/AVC standard (Fig. 4). It comprises a 4-layer hierarchical search. The first layer applies a combination of a telescopic search (TS) and a direct search (DS) with 2-pel precision. While it can realize a wide search range, $D \times S$, (where D is the distance from the current picture to the farthest reference picture, and S is a range of one-step of a telescopic), the TS may mislead to a local minimum motion vector (MV) in case of, for example, camera flash scenes. Then, several direct searches between the current picture and the reference pictures are added for reinforcement. The range of each DS is also S, centering some points such as $(0, 0)$, predicted motion vectors (PMVs), or the MVs of the previous coding in case of transcoding or tandem coding.

The second to fourth layers are full searches with ± 1-pel, ± 0.5-pel, and ± 0.25-pel ranges centering the MVs obtained from the upper layers. While the first and second layers search MVs of four 8×8 blocks in a macroblock (MB), the third and fourth layers evaluate MVs of all block sizes supported (8×8, 8×16, 16×8, and 16×16) obtained from the 8×8 MVs.

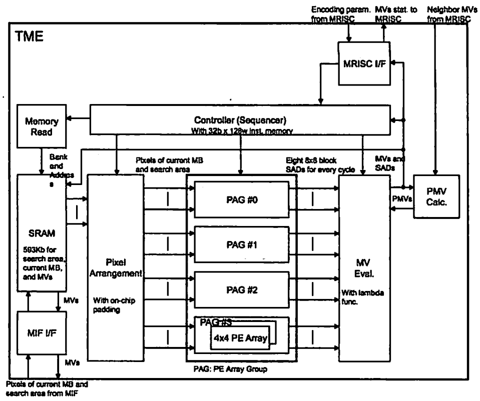

図 5　TME の構成.
Fig. 5　Block diagram of the TME.

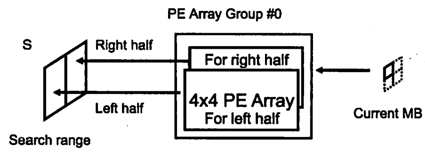

図 6　PE アレイ・グループにおける並列化.
Fig. 6　Parallel PE arrays in a PE-array group.

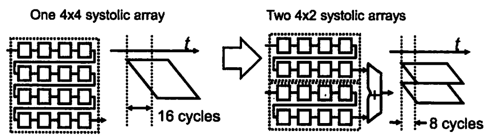

図 7　PE アレイ内における並列化.
Fig. 7　Parallel systolic arrays in a PE array.

3.2 TME

The architecture of the TME, which can executre the first layer search, is shown in Fig. 5. Four PE Array Groups (PAGs) corresponded with four 8×8 blocks in an MB work in parallel. Two types of parallelism are additionally introduced in each PAG to realize a wide search range. First, in order to make S double, each PAG has twin PE arrays that search left and right halves of the search range. Secondly, a 4×4 systolic array (SA) in a PE array is divided into two 4×2 SAs. It takes 16 cycles from the start of one-step search for a 4×4 SA to output the first SAD. The next step search cannot start until the previous search results are fixed in the TS. Two 4×2 SAs can make the start-up cycles half, and increase D from 7 to 9.

3.3 SME

Fig. 8 shows the SME that performs the 0.5- and 0.25-pel

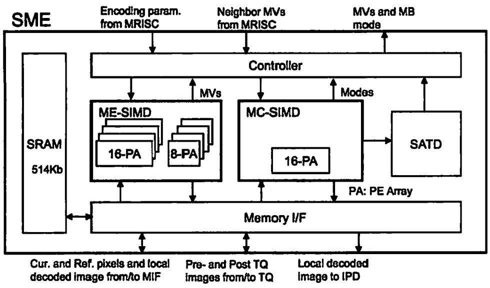

図 8 SME の構成.
Fig. 8 Block diagram of the SME.

ME and MC [8]. It consists of two kinds of SIMD processors: one for ME and the other for MC. The ME-SIMD has four 16-PE arrays and four 8-PE arrays to calculate SADs of variable block size, 8×8, 8×16, 16×8, and 16×16.

The MC-SIMD consists of a 16-PE array and executes various operations with flexible datapath, bi-directional prediction, temporal-/spatial-direct modes, inter/intra decisions, and so on. Fig. 9 shows examples of the flexible datapath of the MC-SIMD. A 4:2:2 chroma format can be supported by changing the MC-SIMD's programs. All executions of SME for an MB are carefully scheduled in consideration of the complex data dependency (Fig. 10).

4. Implementation

A microphotograph of the SARA/E is shown in Fig. 11, and the chip specifications are summarized in Table 1. It was successfully fabricated in a 90 nm technology. It integrates 140 million transistors. The chip can encode D1 (720×480, 30 fps) in real time. With multi-chip configurations on a post-card size module (Fig. 12), it can encode full HDTV (1920×1080, 30 fps). The SARA/E supports High422 profile (8bit only) of H.264/AVC and 422 profile of MPEG-2. All of coding structures, field, frame, PAFF, MBAFF, are supported. Search range of ME is $-217.75/+199.75$ horizontally, and $-109.75/+145.75$ vertically. Almost all ME coding tools are supported, such as multiple reference frames, variable block size, weighted prediction, and spatial/temporal direct mode. High coverage of the H.264/AVC encoding tools in the SARA/E means the high potential of the chips.

5. Evaluations

Fig. 13 shows image quality comparison between the SARA/E and the JM 12.4, a reference software. For fast moving scenes, like scene id 2, 6, and 7, our chip has 1.2 to 1.7 dB gain compared to the JM 12.4. This is because the ME/MC engines can find better MVs. The SARA/E also has advantage of an average image quality with 0.3 dB.

Fig. 14 shows automatic weighted prediction (WP) in fade scenes using the ACC scheme. In these graphs, X-axis shows the number of pictures, and Y-axis shows average values of luminance in decoded pictures. Without WP, the lines of fade-in and fade-out are jaggy. It degrades subjective evaluations. With automatic WP, the lines of fade-in and fade-out become smooth. Thus, the SARA/E can encode a variety of scenes efficiently.

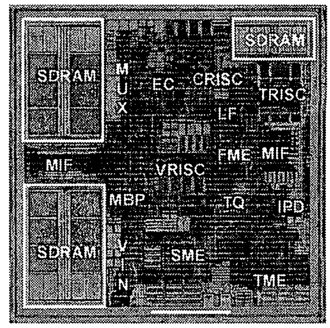

図 11 SARA/E チップ写真.
Fig. 11 Microphotograph of the SARA/E.

表 1 SARA/E 諸元.
Table 1 Specification of the SARA/E.

Technology		90-nm
Number of transistors		140 million transistors
Clock frequency		200MHz (max)
Supply voltage		Core: 1.2V / MobileDDR: 1.8V / eDRAM: 2.5V / I/O: 3.3V
Power consumption		3.0W
Package		625-pin FCBGA (21mm x 21mm)
Memory		eDRAM: 72Mbit, external: 512Mbit (32-bit width) MobileDDR
Video	Profile	H.264: Main / High / High422 (8bit only), MPEG2: Main / 422P
	Level	H.264: 3.0 / 4.0 / 4.1, MPEG2: ML / H14L / HL
	Resolution and video rate	Single chip: 720 x 480 at up to 30fps Multiple chip: 1920 x 1080 at up to 30fps
	Coding structure	Field, Frame, PAFF, MBAFF
	Motion estimation	Search range: -217.75 / +199.75 (H), -109.75 / +145.75 (V) Max. number of reference frames: 4 Supported block size: 8x8, 8x16, 16x8, 16x16 Weighted prediction mode: explicit Direct mode: spatial, temporal
	Pre-processing	Adaptive temporal and spatial filter Macroblock-based feature extraction
	Transcoding	Using mode information or our original information

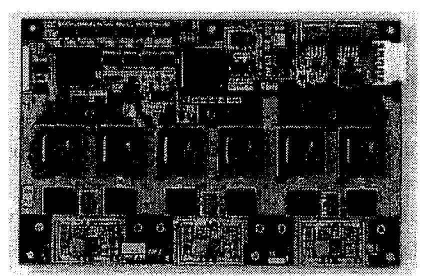

図 12 SARA/E HD モジュール.
Fig. 12 SARA/E HD module.

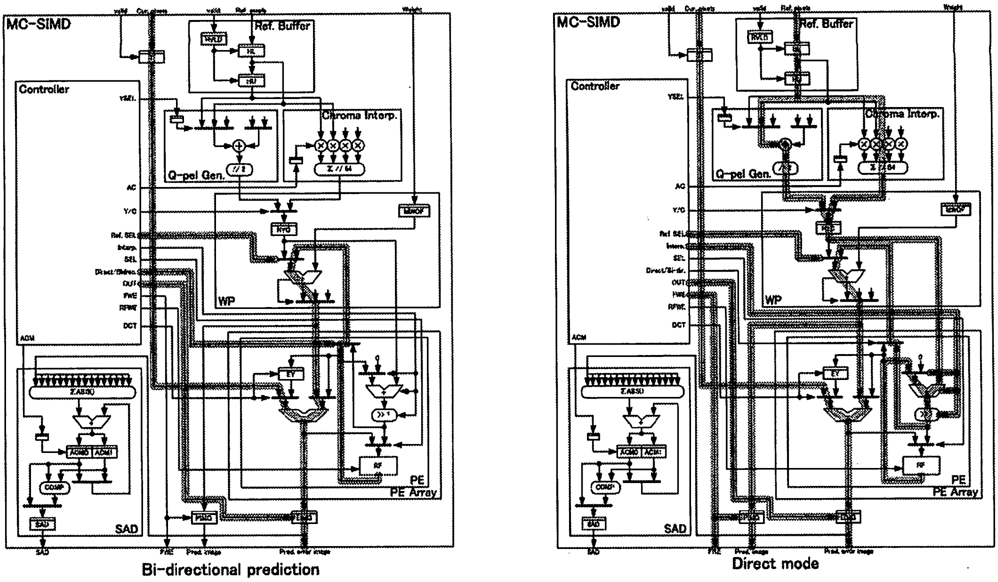

図 9 MC-SIMD のデータパス.
Fig. 9 Flexible datapath of MC-SIMD.

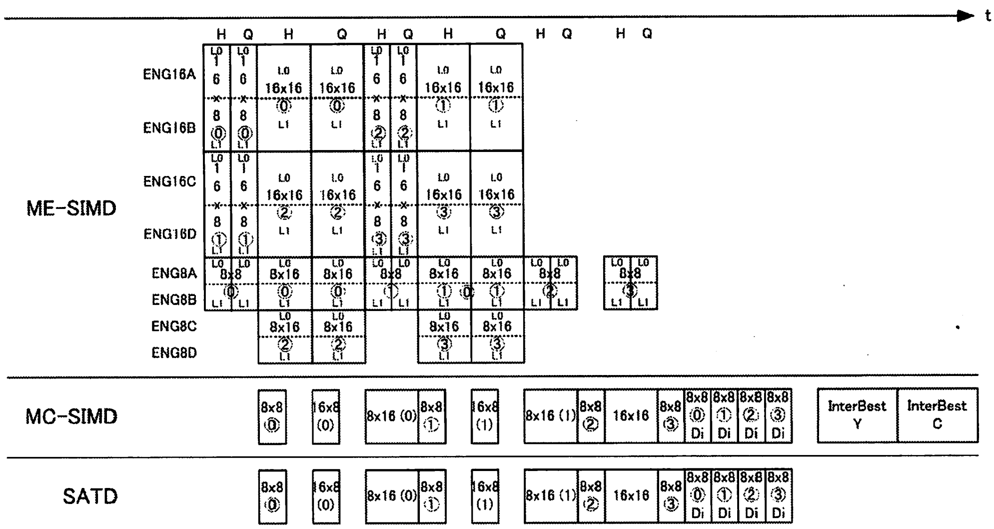

図 10 SME パイプライン・スケジュール.
Fig. 10 Pipeline schedule of the SME.

6. Conclusions

We have developed an H.264/AVC and MPEG-2 encoder LSI, SARA/E, for HDTV broadcasting infrastructures. It has pre-analysis engines for the advanced coding control scheme. The ACC scheme can realize automatic weighted prediction, and transcoding with inheriting precoding modes. It also has powerful ME/MC blocks that can realize wide search range, multiple reference frame, variable block size, and weighted prediction.

The SARA/E chips are compactly mounted on a postcard size HD module, which enables us to build up various CODEC equipments. Besides a simple (1-passed) encoder (Fig. 15), a transcoder with a MPEG-2 decoder can be used for re-transmission services of digital terrestrial broadcasting over IP network such as our NGN, a tandem encoder with

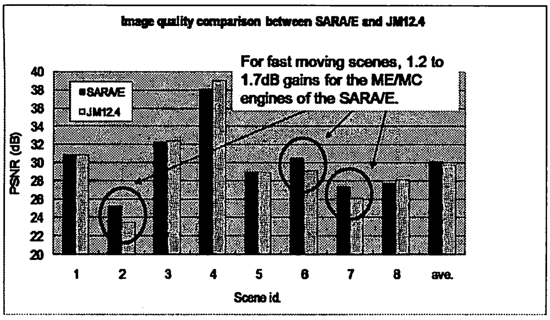

図 13 SARA/E と JM との画質比較.

Fig. 13 Image quality comparison between the SARA/E and the JM 12.4.

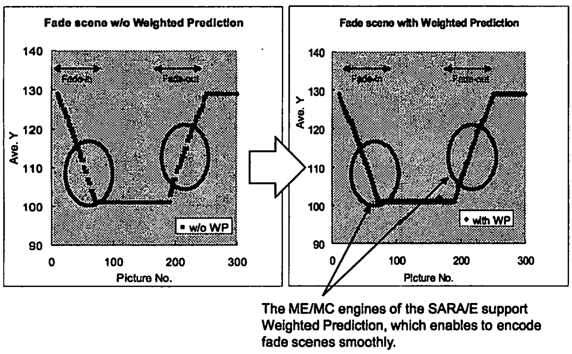

図 14 フェード・シーンでの ACC の効果.

Fig. 14 Fade scene using ACC.

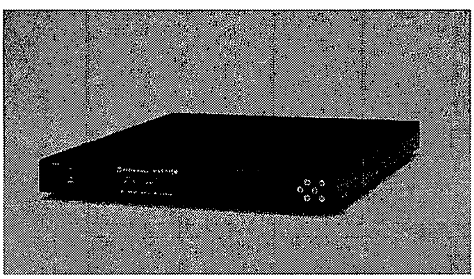

図 15 SARA/E を用いたエンコーダ装置.

Fig. 15 Encoder equipment using the SARA/E.

two HD modules can be developed as a high compression encoder.

The SARA/E will be a key device for implementing various professional H.264/MPEG-2 applications for future broadcasting infrastructures.

文 献

[1] "Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding" (2003). ISO/IEC 14496-10:2003.

[2] Y.-W. Huang, T.-C. Chen, C.-H. Tsai, C.-Y. Chen, T.-W. Chen, C.-S. Chen, C.-F. Shen, S.-Y. Ma, T.-C. Wang, B.-Y. Hsieh, H.-C. Fang and L.-G. Chen: "A 1.3TOPS H.264/AVC single-chip encoder for HDTV applications", ISSCC Digest of Technical Papers, Vol. 1, pp. 128–129,588 (2005).

[3] H.-C. Chang, J.-W. Chen, C.-L. Su, Y.-C. Yang, Y. Li, C.-H. Chang, Z.-M. Chen, W.-S. Yang, C.-C. Lin, C.-W. Chen, J.-S. Wang and J.-I. Quo: "A 7mW-to-183mW dynamic quallity-scalable H.264 video encoder chip", ISSCC Digest of Technical Papers, pp. 280–281,603 (2007).

[4] H. Mizosoe, D. Yoshida and T. Nakamura: "A single chip H.264/AVC HDTV encoder/decoder/transcoder system LSI", ICCE Digest of Technical Papers, pp. 1–2 (2007).

[5] K. Nitta, M. Ikeda, H. Iwasaki, T. Onishi, T. Sano, A. Sagata, Y. Nakajima, M. Inamori, T. Yoshitome, H. Matsuda, R. Tanida, A. Shimizu, K. Nakamura and J. Naganuma: "An H.264/AVC High422 profile and MPEG-2 422 profile encoder LSI for HDTV broadcasting infrastructures", IEEE Symposium on VLSI Circuits, pp. 106–107 (2008).

[6] H. Iwasaki, J. Naganuma, K. Nitta, K. Nakamura, T. Yoshitome, M. Ogura, Y. Nakajima, Y. Tashiro, T. Onishi, M. Ikeda, T. Minami, M. Endo and Y. Yashima: "Single-chip MPEG-2 422P@HL CODEC LSI with multi-chip configuration for large scale processing beyond HDTV level", IEEE Transaction on VLSI Systems, **15**, pp. 1055–1059 (2007).

[7] K. Suguri, T. Minami, H. Matsuda, R. Kusaba, T. Kondo, R. Kasai, T. Watanabe, H. Sato, N. Shibata, Y. Tashiro, T. Izuoka, A. Shimizu and H. Kotera: "A real-time motion estimation and compensation LSI with wide search range for MPEG2 video encoding", IEEE Journal of Solid-State Circuits, **31**, pp. 1733–1741 (1996).

[8] T. Onishi, T. Sano, K. Nitta, M. Ikeda and J. Naganuma: "Multi-reference and multi-block-size motion estimation with flexible mode selection for professional 4:2:2 H.264/AVC encoder LSI", IEEE International Symposium on Circuits and Systems (ISCAS2008), pp. 800–803 (2008).