A Low-Power H.263 Video CoDec Core Dedicated to Mobile Computing

Morgan Hirosuke MIKI,

Gen FUJITA,

Takeshi KOBAYASHI,

Takao ONOYE, and

Isao SHIRAKAWA

Dept. Information Systems Engineering

Osaka University

2-1 Yamada-Oka, Suita, Osaka, 565 Japan

Phone: +81(6)879-7807, FAX: +81(6)875-5902

e-mail: {miki, fujita, kobayash, onoe,sirakawa@ise.eng.osaka-u.ac.jp}

Abstract

A number of novel VLSI architectures are devised for an H.263 video codec core in terms of low bitrate visual communication. The potential of the practicability for mobile computing has been extremely explored by attempting not only to minimize the total chip area but also to reduce the power consumption to such an extent that the operation frequency can be slowed down to 15MHz. The whole encoding and decoding facilities have been integrated in the die area of 7.66 mm^2 by means of a 0.35m CMOS technology, with the dissipation of 146.60 mW from a single 3.3V supply.

Keywords

VLSI, video codec, low bitrate, H.263, low power

1   INTRODUCTION

The H.324(ITU-T, 1995) international standard specifies the low bitrate audiovisual communication based on PSTN (Public Switched Telephone Network). The H.263(ITU-T, 1996) standard is a video version of this H.324, which is to compress the moving picture components of audiovisual services at low bitrates. Actually, by means of H.263, QCIF (176 × 144) 10 fps (frames/sec) pictures can be coded at V.34 (28.8Kbps), and moreover it should be added that at such a low bitrate the H.263 coding efficiency is superior to any of those of H.261(ITU-T, 1993) and MPEG1(ISO/IEC, 1993) . Thus various applications of this H.263 standard are to be realized extensively in mobile computing, wireless multimedia communication, etc. In particular, portable multimedia facilities in the wireless environment can be regarded as the enormous potentialities of multimedia communications.

The coding/encoding process of the H.263 standard may be implemented with the use of any of those multimedia enhanced DSPs(Brinthaupt, 1996)(Golston, 1996)(Okamoto, 1996), which have been developed specifically for H.261 or MPEG1. However, in terms of the mobile and portable use, there still remains much room for reducing both the power consumption and chip area of the codec core. Thus there arises the big issue of how to develop the H.263 specific architectures of VLSI implementation especially for mobile computing.

The present paper describes a number of VLSI architectures sophisticated for implementing an H.263 video codec dedicated principally to mobile computing. The main distinctive features of this codec consists in a large reduction not only in the total chip area but also in the power consumption to such an extent as to slow down the operation frequency to 15MHz. All of encoding/decoding facilities have been integrated in the area of 7.66 mm^2 by a 0.35m triple-metal CMOS technology, with the total dissipation of
146.60 mW from a single 3.3V supply.

2   H.263 VIDEO CODING ALGORITHM

The main encoding/decoding process of H.263 is the so-called MC-DCT coding, as shown in Figure 1, which is executed in the same manner as H.261 and MPEG1. The distinctive features of the H.263 standard lie in simple syntax, half-pel prediction, block-level motion estimation (advanced prediction mode), paired coding of the P-frame and the B-frame (i.e. the PB-frame mode), motion detection for the outside of ̈ frame (i.e. the unrestricted vector mode), SAC (syntax-based arithmetic coding mode), and so on.

As can be seen from Figure 1, the picture coding can be achieved with the use of several functional units; [Motion Estimator (ME)], [Discrete Cosine Transformer (DCT)], [Quantizer (Q)], and [Variable Length Coder (VLC)]/[Syntax-based Arithmetic Coder (SAC)]. The bitstream decoding can be performed with the use of

another set of functional units; [Variable Length Decoder (VLD)]/[Syntax-based Arithmetic Decoder (SAD)], [Inverse Quantizer (IQ)], [Inverse Discrete Cosine Transformer (IDCT)], and [Motion Compensator (MC)]. The typical picture encoding format considered here is QCIF (176 × 144) 10-15fps (frames/sec) at 28.8Kbps.
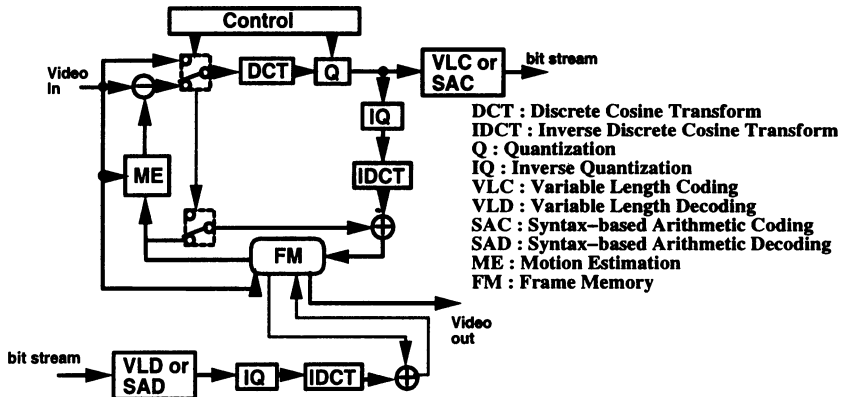

DCT : Discrete Cosine Transform
IDCT : Inverse Discrete Cosine Transform
Q : Quantization
IQ : Inverse Quantization
VLC : Variable Length Coding
VLD : Variable Length Decoding
SAC : Syntax–based Arithmetic Coding
SAD : Syntax–based Arithmetic Decoding
ME : Motion Estimation
FM : Frame Memory

Figure 1 H.263 encoding/decoding process

3   VLSI ARCHTECTURE
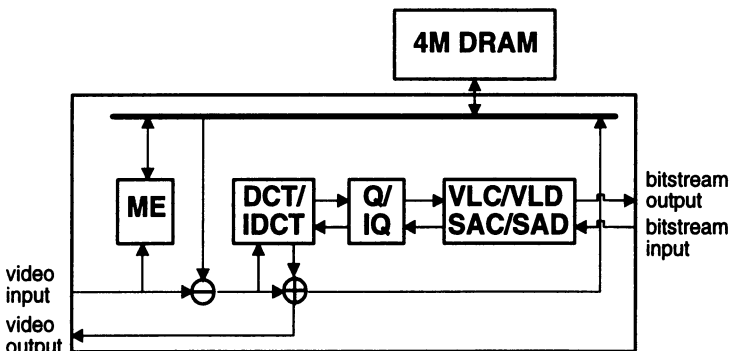
3.1 Organization

Figure 2 Organization of H.263 codec core

Seeing that this H.263 video codec core is intended for the single chip implementation of a realtime H.324 audiovisual codec, innovations should be devised not only in reducing the area occupancy and the power dissipation but also in refining on the external memory size and the memory accessing bandwidth. The overall organization of our codec core is summarized in Figure 2, which is composed of a number of specific functional units. The main factor to achieve a high throughput at a low operation frequency can be attributed to the mechanism

that the I/O and processing conflicts can be mitigated at each stage of the codec. The detailed architecture of each functional unit is outlined in what follows.

3.2 ME Core

As to the so-called block-matching algorithm for ME (Motion Estimator), a number of authors(Chen, 1995)(Hayashi, 1995)(Onoye, 1995)(Chan, 1995)(Kim, 1995) have attempted to reduce the computational costs of the *full-search*, which is to detect a motion vector exhausitively within a search range by reference to MADs (Mean Absolute Differences). A compact ME core(Fujita, 1997) has been attained for H.263 by means of a sophisticated *macroblock clustering* algorithm(Onoye, 1996), which has the following features;

1.   high quality vectors,
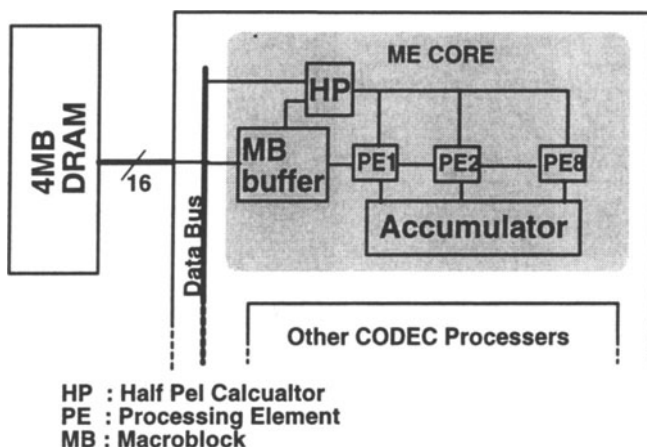2.   low computational costs, and
3.   VLSI implementation capability.

HP : Half Pel Calcualtor
PE : Processing Element
MB : Macroblock

Figure 3 Organization of ME core The H.263 standard supports the PB-frame

The organization of the ME core is illustrated in Figure 3, which consists of a one-dimensional PE (Processing Element) array, an accumulator for calculating macroblock vectors and block vectors, a macroblock buffer for the bi-directional prediction, and a half-pel calculator.

Figure 4 indicates a block diagram of a PE, which adopts 8-bit and 12-bit datapath circuits. The reference pixel is to be broadcast to all PEs, and the prediction pixel is to be propagated from PE to PE. A PE outputs an MAD of 8 pixels at every 8 cycles.
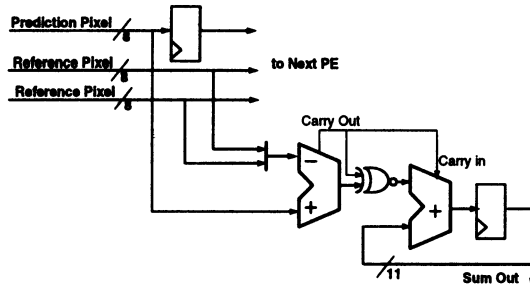
Figure 4 Block diagram of PE

In addition to the normal macroblock prediction, the H.263 standard supports the advanced prediction mode, which is to detect motions of four blocks in a macroblock. In other words, as outlined in Figure 5, a macroblock can have either one macroblock vector or four block vectors. To cope with this, the organization of the accumulator is devised as illustrated in Figure 6. The MADs for macroblock and four blocks are calculated simultaneously by accumulating 8 pixels' MADs output from the PEs.
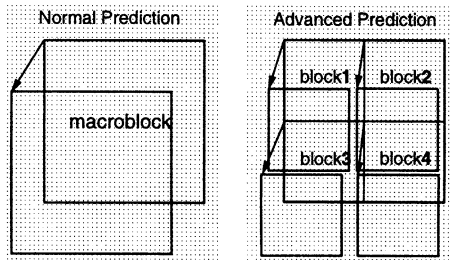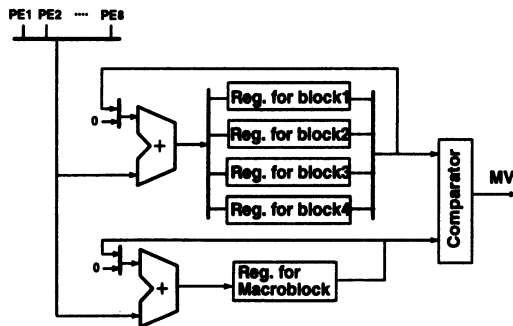

Figure 5 Advanced prediction mode

Figure 6 Block diagram of accumulator

The H.263 standard supports the PB-frame mode so that two frames (P-frame and B-frame) can be coded as one unit, and the ME core seeks the vectors for a pair of

macroblocks of these two frames simultaneously. Apart from MPEG, the motion estimation for the B-frame of H.263 requires the concurrent reference of two frames, since only one vector per marcoblock is used for the bi-directional prediction as illustrated in

Figure 7. Therefore, the half-pel calculator determines the average of forward and backward reference pixel data, and feeds them to the PE array. The former reference pixel data are read from the external memory, and the later from the macroblock buffer.
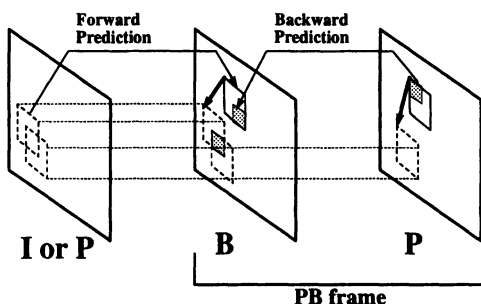

Figure 7 Bi-directional prediction of H.263

3.2 DCT/IDCT Core

The DCT has been successfully employed so far in a variety of algorithms(ITU-T, 1993)(ITU-T, 1994)(ITU-T, 1993)(ISO/IEC, 1993) for the image compression to reduce the spational redundancy of picture sequence.

The computational costs of the H.263 codec is lower than those of the MPEG1/2 cores. The DCT/IDCT architectures, which have been developed for MPEG1/2(Matsui, 1994)(Uramoto, 1992)(Masaki, 1995), should not be employed for H.263 from the view point of hardware cost, and therefore in what follows a novel specific architecture is proposed for H.263.

For implementing DCT/IDCT core the Chen's algorithm(Chen, 1977) (butterfly computation) is widely used in conjunction with a distributed arithmetic(Peled, 1974). The Chen's algorithm can reduce the number of multiplications in DCT/IDCT by half. According to this algorithm, the 8×1 DCT and 8×1 IDCT are calculated by means of the following equations,

$$
\text{DCT:} \quad
\begin{bmatrix} X_0 \\ X_2 \\ X_4 \\ X_6 \end{bmatrix}
=
\begin{bmatrix} A & A & A & A \\ B & C & -C & -B \\ A & -A & -A & A \\ C & -B & B & -C \end{bmatrix}
\begin{bmatrix} x_0 + x_7 \\ x_1 + x_6 \\ x_2 + x_5 \\ x_3 + x_4 \end{bmatrix}
\tag{1}
$$

$$
\begin{bmatrix} X_1 \\ X_3 \\ X_5 \\ X_7 \end{bmatrix} = \begin{bmatrix} D & E & F & G \\ E & -G & -D & -F \\ F & -D & -G & E \\ G & -F & E & -D \end{bmatrix} \begin{bmatrix} x_0 - x_7 \\ x_1 - x_6 \\ x_2 - x_5 \\ x_3 - x_4 \end{bmatrix} \tag{2}
$$

IDCT:
$$
\begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} A & B & A & C \\ A & C & -A & -B \\ A & -C & -A & B \\ A & -B & A & -C \end{bmatrix} \begin{bmatrix} X_0 \\ X_2 \\ X_4 \\ X_6 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} D & E & F & G \\ E & -G & -D & -F \\ F & -D & G & E \\ G & -F & E & -D \end{bmatrix} \begin{bmatrix} X_1 \\ X_3 \\ X_5 \\ X_7 \end{bmatrix} \tag{3}
$$

$$
\begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} A & B & A & C \\ A & C & -A & -B \\ A & -C & -A & B \\ A & -B & A & -C \end{bmatrix} \begin{bmatrix} X_0 \\ X_2 \\ X_4 \\ X_6 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} D & E & F & G \\ E & -G & -D & -F \\ F & -D & G & E \\ G & -F & E & -D \end{bmatrix} \begin{bmatrix} X_1 \\ X_3 \\ X_5 \\ X_7 \end{bmatrix} \tag{4}
$$

$$
A = \cos\frac{\pi}{4}, \quad B = \cos\frac{\pi}{8}, \quad C = \sin\frac{\pi}{8}, \quad D = \cos\frac{\pi}{16},
$$

$$
E = \cos\frac{3\pi}{16}, \quad F = \sin\frac{3\pi}{16}, \quad G = \sin\frac{\pi}{16},
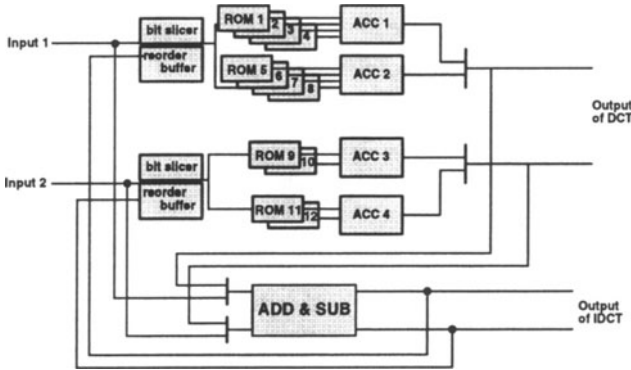$$

Figure 8 DCT/IDCT core by distributed arithmetic

By means of the distributed arithmetic, as illustrated in Figure 8, each multiply accumulation can be executed with the use of accumulators and those ROMs which contain tables of products calculated in advance. However, this scheme requires additionally *bit slicer* and *reorder buffer*, and these units as well as ROMs occupy considerable area in the case of H.263. That is, to implement H.263 DCT/IDCT core, this overhead must be a serious obstacle.

On the contrary, as illustrated in Figure 9, our DCT/IDCT is devised dedicatedly for H.263, where ROMs **ROM A ~ ROM G** and accumulators **ACC 1 ~ ACC 4** calculate each multiplication of equations (1)-(4) (i.e. $A(x_0 + x_7)$, $A(x_1 + x_6)$, ..., AX_0, BX_2, ..., etc.), 4 bits at a time. As a result, the reorder buffer can be removed, and the number of ROMs can be reduced without degrading a performance.
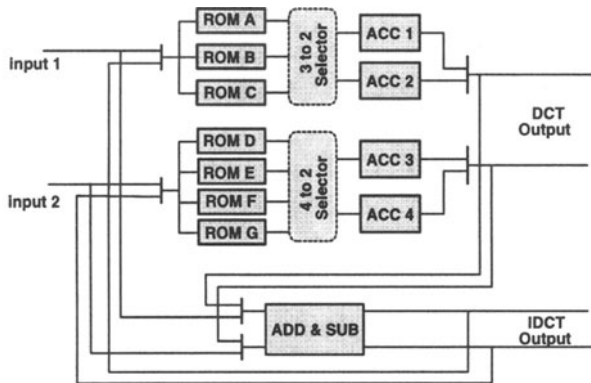
Figure 9 Block diagram of proposed DCT/IDCT core

3.4 Q/IQ Core

The operations of Q and IQ are simply divisions and multiplicatons. To reduce the area occupancy, our Q is implemented by means of a 2-bit sequential non-restorning divider and IQ by means of a radix-4 sequential Booth's multiplier without parallel/array facilities, as illustrated in Figure 10. Both of these Q and IQ can output the division/multiplication result at every 4 cycles.
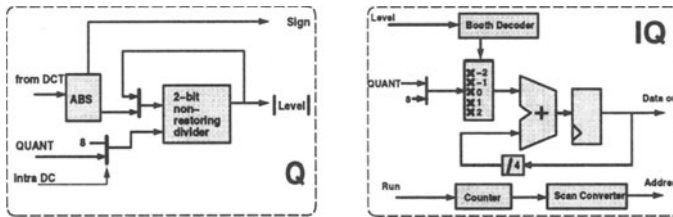

Figure 10 Block diagram of Q/IQ core

3.5 VLC/VLD and SAC/SAD Core

In addition to VLC, the H.263 standard supports SAC, which is based on arithmetic operations and table index search. In order to achieve high coding efficiency, either of those two encoding modes is choosen picture by picture.

Figure 11 shows a block diagram of the coding core. In both of two coding modes, a set of *run*, *level*, and *last* is indicated by an index, and then the index is coded to bitstream. Thus, it turns out that the *index generator* can be shared by different coding modes. As for the VLC table, the compressing mechanism proposed by (Tanaka, 1995) is employed to reduce the table size. The *arithmetic unit* calculates arithmetic operations required for SAC and SAD.
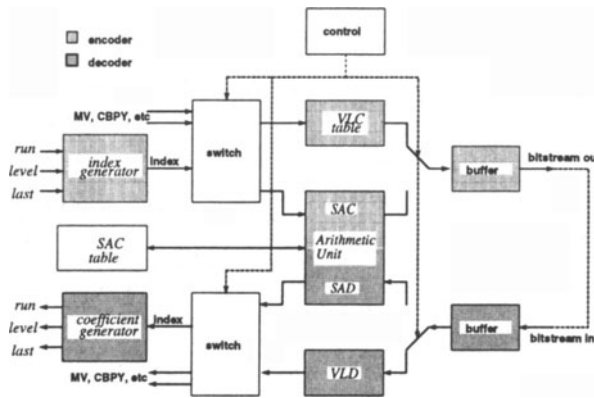
Figure 11 Block diagram of VLC/VLD and SAC/SAD core

4   IMPLEMENTATION RESULTS

A set of sophisticated architectures so far outlined are implemented through the use of ASIC design system *COMPASS Design Tools ver. 9*. This system allows the top down design of high-performance ASIC from the hybrid design entry of circuit/datapath schematic and HDL. For one implementation, the high-speed datapaths are used in the ME and the DCT/IDCT core, which require considerable computation. The standard cell logic blocks for controllers and ROM tables are synthesised from VHDL descriptions. It should be added in orther that the operation frequency slows down to 15MHz, the total power dissipation is reduced to 146.60 mW, and hence the core can be of portable use. Table 1 indicates the main chip features of the codec core, and Figure 12 shows the layout patterns obtained by a 0.35 *m* triple-metal technology.

Table 1 Main chip feature of codec core

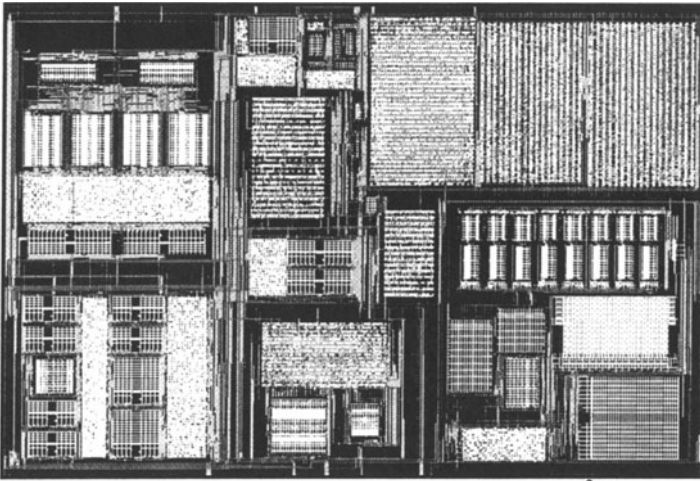|	
Technology	0.35*m* CMOS triple-level Al.
Chip size	3.39 *mm* × 2.26 *mm*
Transistors	187,266
Clock frequency	15.0 MHz
Power Dissipation	146.60 mW (3.3 V, 15.0 MHz)
Support picture	QCIF(176 × 144), sub-QCIF(128 × 96) 10fps
Encoding Options	advanced prediction mode, PB-frame mode, unrestricted vector mode, syntax-based arithmetic coding mode

Figure 12 Layout pattern of H.263 codec core ($3.39 \times 2.26 \ mm^2$)

5   CONCLUSION

This paper has outlined a sophisticated set of VLSI architectures for an H.263 codec core, dedicated to mobile computing. Specifically, the ME core can treat various encoding options, and multipliers and dividers at a low operational frequency are employed in the Q/IQ core and the VLC/VLD and SAC/SAD core.

Development is continuing on an integrated set of architectures for the single chip implementation of H.324 audiovisual communication.

6   REFERENCES

Brinthaupt, D., Knoblock, J., Othmer, J., Petryna, B., Uyttendaele, M. (1996) A programmable audio/video processor for H.320, H.324, and MPEG, in *IEEE ISSCC Digest of Technical Papers*, 292-293.

Chen, M. C. and Willson Jr, A. N. (1995) A high accuracy predictive logarithmic motion estimation algorithm for video coding, in *Proc. IEEE Int'l Symp. Circuits and Systems*, 617-620.

Chen, W. H., Smith, C. H., and Fralick, S. C. (1977) A fast computational algorithm for the discrete cosine transform, *IEEE Trans. Communications*, **COM-25**, 9, 1004-1009.

Chan, Y.-L. ,and Siu (1995), W.-C. A new block motion vector estimation using adaptive pixel decimation, in *Proc. IEEE Int'l Conf. Acoustics, Speech, and Signal Processing*, 2257-2260.

Fujita, G., Onoye, T., and Shirakawa, I. (1997) A new motion estimation core dedicated to H.263 video coding, in *Proc. IEEE Int'l Symp. Circuits and Systems*, to apper.

Golston, J. (1996) Single-chip H.324 videoconferencing, *IEEE Micro*, **16**, 4, 21-33.

Hayashi, N., Kitsuki, T. , Tamitani, I. , Honma, H., Ooi, Y., Miyazaki, T., and Oobuchi, K. (1995) A bidirectional motion compensation LSI with a compact motion estimator, *IEICE Trans. Electron*, **E78-C**, 12, 1682-1690.

ITU-T Rec.H.261 (Mar. 1993) Video codec for audiovisual services at $p \times 64$ kbits, International Standard.

ITU-T Rec. H. 262, ISO/IEC 13818-2 (1994) Generic coding of moving pictures and associated audio, Draft International Standard.

ITU-T Rec. H.324 (1995) Terminal for low bitrate multimedia communication, Draft International Standard.

ITU-T Rec. H.263 (1996) Video coding for low bitrate communication, Draft International Standard.

ISO/IEC 11172-2 (1993) Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s, International Standard.

Kim, Y., Rim, C.S. and Min (1995), B. A block matching algorithm with 16:1 subsampling and its hardware design, in *Proc. IEEE Int'l Symp. Circuits and Systems*, 613-616.

Matsui, Y, Hara, H., Uetani, Y., Kim, L.-S., Nagamatsu, T., Watanabe, Y., Chiba, A., Matsuda, K., and Sakurai, V (1994) A 200MHz 13 mm^2 2-D DCT macrocell using sense-amplifying pipeline flip-flop scheme, *IEEE J. Solid-State Circuits*, **29**, 12, 1482-1490.

Masaki, T., Morimoto, Y., Onoye, T., and Shirakawa, I. (1995) VLSI implementation of inverse discrete cosine transformer and motion compensator for MPEG2 HDTV video decoding, *IEEE Trans. Circuits and Systems for Video Technology*, **5**, 5, 387-395.

Okamoto, K., Jinbo, T., Araki, T., Iizuka, Y., Nakajima, H., Takahata, M., Inoue, H., Kurohmaru, S., Yonezawa, T. and Aono, K. (1996) A DSP for DCT-based and wavelet-based video CODEC's for consumer applications, in *Proc. IEEE Custom Integrated Circuits Conference*, 359-362.

Onoye, T., Takatsu, M., Fujita, G., and Shirakawa, I. (1995) A VLSI-suited motion estimation algorithm based on macroblock clustering, *IEICE Tech. Report*, CAS95-43 (in Japanese).

Onoye, T., Fujita, G., Takatsu, M., Shirakawa, I., and Yamai, N. (1996) Single chip implementation of motion estimator dedicated to MPEG2 MP@HL, *IEICE Trans. Fundamentals*, **E79-A**, 8, 1210-1216.

Peled, A. and Liu, B. (1974) A new hardware realization of digital filters, *IEEE Trans. Accoust., Speech, Signal Processing*, **ASSP-22**, 6, 456-462.

Tanaka, E. and Enomoto, T. (1995) 70mW variable length codec for MPEG2, in *Proc. IEICE General Conference*, **C-577**.

Uramoto, S., Inoue, Y., Takabatake, A., Takeda, V, Yamashita,V, Terane, H., and Yoshimoto, M. (1992) A 100MHz 2-D discrete cosine transform core processor, *IEEE J. Solid-State Circuits*, **27**, 4, 492-499.

7 BIOGRAPHY

Morgan Hirosuke Miki was born in Atibaia-SP, Brazil, on August 3, 1971. He received B.E degree in electronic engineering from University of São Paulo, São Paulo, Brazil, in 1995. He is now working toward the M.E degree in information systems engineering, Osaka University. His research interest include Computer-Aided Design of VLSI Circuits.

Gen Fujita was born in Kobe, Japan, on October 20, 1971. He received B.E. and M.E. degree in information systems engineering from Osaka University, Osaka, Japan, in 1994 and 1996 respectively. He is now working toward the Ph.D. degree in the Department of Information Systems Engineering, Osaka University. His research interests include computer-aided design of VLSI Circuits.

Takeshi Kobayashi was born in Takarazuka, Japan, on January 16, 1974. He received B.E degree in information systems engineering from Osaka University, Osaka, Japan, in 1996. He is currently with KDD Co., Ltd.

Takao Onoye was born in Kobe, Japan on May 9, 1968. He received B.E. and M. E. degrees in electronic engineering from Osaka University, Osaka, Japan, in 1991 and 1993 respectively. He is currently a research assistant of the Department of Information Systems Engineering, Osaka University. His research interest include computer-aided design and Implementation of application specific VLSI's, especially in the field of image generation, moving picture image processing(MPEG2 CODEC), etc.

Isao Shirakawa was born in Toyama Pref., Japan, on September 12, 1939. He received the B.E., M.E and Ph.D. degrees, all in electronic engineering, from Osaka University, Osaka, Japan, in 1963, 1965, and 1968, respectively. He joined the Department of Electronic Engineering at the same University in 1968 as a research assistant, where he was promoted to Associate Professor in 1973 and then to Professor in 1987, and is now Professor in Department of Information Systems Engineering. Meanwhile, he was with the Electronic Research Lab., University of California, Berkeley, as a Visiting Scholar in 1974-1975. During 1995-1996 he was a vice-president of the IEEE CAS society. He has been engaged in education and research mainly on basic circuit theory, applied graph theory, VLSI synthesis.