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

Morgan H. Miki, Gen Fujita, Takao Onoye, and Isao Shirakawa
Dept. Information Systems Engineering, Osaka University
Phone: +81(6)879-7808, FAX: +81(6)875-5902
e-mail: {mikifuji,taoonoe,sirakawa}@ise.eng.osaka-u.ac.jp

ABSTRACT

A low-power H.263 video codec core dedicated to low bitrate visual communication is described. A number of sophisticated architectures have been devised 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. As a result, the whole encoding and decoding facilities of an H.263 video codec core have been integrated in the die area of $6.54 mm^2$ by means of a $0.35 \mu m$ CMOS technology, with the dissipation of 146.60mW from a single 3.3V supply.

1 INTRODUCTION

The H.324[1] international standard specifies the low bitrate audiovisual communication based on PSTN (Public Switched Telephone Network). The H.263[2] 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[3] and MPEG1[4]. 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 encoding/decoding process of the H.263 standard may be implemented with the use of those multimedia-enhanced DSPs[5, 6, 7, 8], which have been developed specifically for H.261 and/or MPEG1. In this case, however, the power consumption can hardly be held down any more, since those DPSs should operate at a high frequency to raise the performance, which forces an H.263 codec core to dissipate 500mW or more. Thus, 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, and hence there arises the big issue of how to develop the H.263 specific architectures of VLSI implementation dedicatedly for portable use.

The present paper describes a number of VLSI architectures sophisticated for implementing an H.263 video codec dedicated to mobile computing. The main feature of this codec consists in a larger 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 of this core have been integrated in the area of $6.54 mm^2$ by a $0.35 \mu m$ triple-metal CMOS technology, with the total dissipation of 146.60mW 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 Fig. 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.

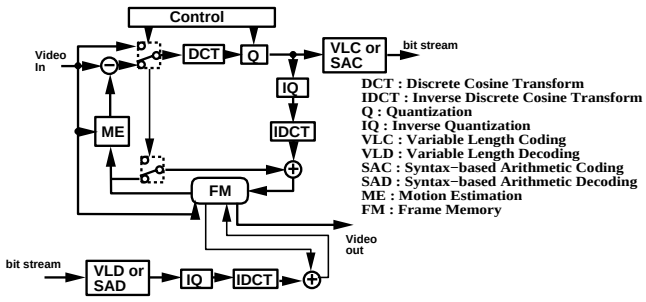


Figure 1: H.263 encoding/decoding process.

As can be seen from Fig. 1, the picture coding can be achieved with the use of several functional units; [Motion Estimator (ME)], [Discret 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.

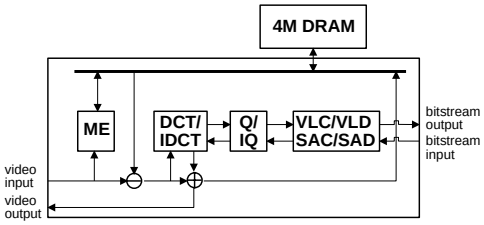


Figure 2: Organization of H.263 codec core.

Seeing that this H.263 video codec core is intended for the single chip implementation of a real-time H.324 audio-visual 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 Fig. 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.

As to the so-called block-matching algorithm for ME (Motion Estimator), a number of authors[9, 10, 11, 12, 13] have attempted to reduce the computational costs of the *full-search*, which is to detect a motion vector exhaustively within a search range by reference to MADs (Mean Absolute Differences). A compact ME core[15] has been attained for H.263 by means of a sophisticated *macroblock clustering* algorithm[14], which has the following features;

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

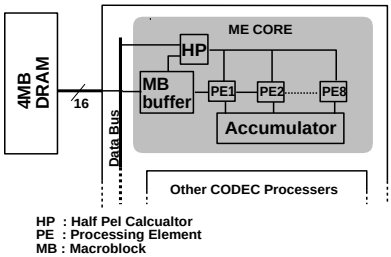


Figure 3: Organization of ME core.

The organization of the ME core is illustrated in Fig. 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.

Fig. 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.

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 Fig. 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 Fig. 6. The MADs for macroblock and four blocks are calculated simultaneously by accumulating 8 pixels' MADs output from the PEs.

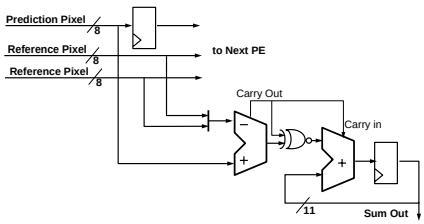


Figure 4: Block diagram of PE

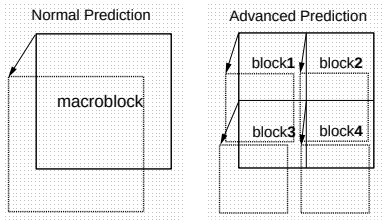


Figure 5: Advanced prediction mode

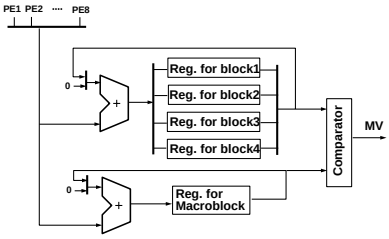
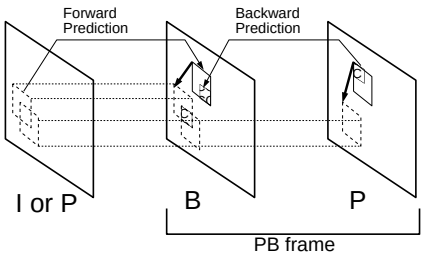


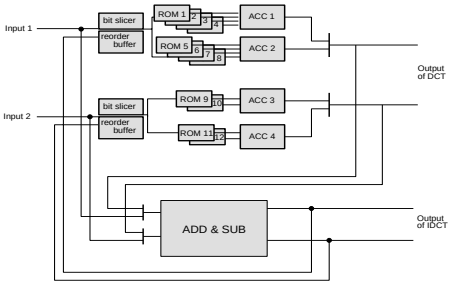
Figure 6: Block diagram of accumulator.

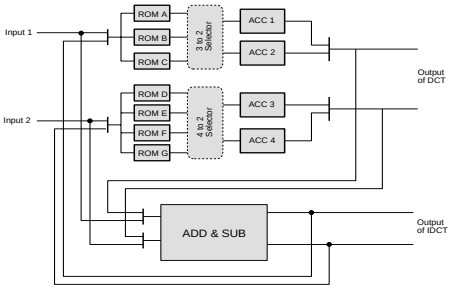
The H.263 standard supports the PB-frame mode so that two frames (P-frame and B-frame) can be coded as one unit, and the ME core seeks the vectors for a pair of macroblocks of these two frames simultaneously. Apart from MPEG, the motion estimation for the B-frame of H.263 requires the concurrent reference of two frames, since only one vector per macroblock is used for the bi-directional prediction as illustrated in Fig. 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.

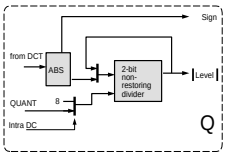
The DCT has been successfully employed so far in a variety of algorithms[2, 3, 4, 16] for the image compression to reduce the spatial redundancy of picture sequence.

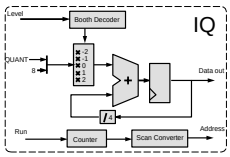
The computational costs of the H.263 codec is lower than those of the MPEG1/2 cores. The DCT/IDCT architectures, which have been developed for MPEG1/2[17, 18, 19], should not be employed for H.263 from the view point of hardware cost, and therefore in what follows a novel specific architecture is proposed for H.263.


Forward
Prediction
Backward
Prediction
I or P
B
P
PB frame


Input 1
Input 2
bit slicer
reorder buffer
bit slicer
reorder buffer
ROM 1
ROM 5
ROM 9
ROM 11
ACC 1
ACC 2
ACC 3
ACC 4
Output of DCT
ADD & SUB
Output of IDCT


Input 1
Input 2
ROM A
ROM B
ROM C
ROM D
ROM E
ROM F
ROM G
3 to 2 Selector
4 to 2 Selector
ACC 1
ACC 2
ACC 3
ACC 4
Output of DCT
ADD & SUB
Output of IDCT


Sign
from DCT
ABS
2-bit non-restoring divider
Level
QUANT
8
Intra DC
Q


Level
Booth Decoder
IQ
QUANT
8
Data out
Run
Counter
Scan Converter
Address

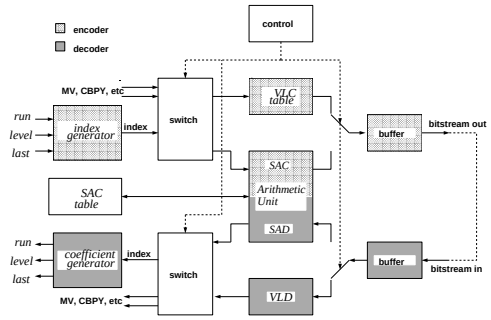


Figure 11: Block diagram of VLC/VLD and SAC/SAD core.

μm CMOS triple-metal technology. It should be added that the operation frequency slows down to 15MHz in order to reduce the total power dissipation to 146.60mW, and hence the core can be of mobile use. Beside this architectural technique, other device level sophisticated low-power technique can allow much more power reduction.

Table 1: Main chip features of codec core.

Technology	0.35 μm CMOS triple-level Al
Chip size	3.3 mm × 1.9 mm
Transistors	187,266
Clock frequency	15.0MHz
Power dissipation	146.60mW (3.3V, 15.0MHz)
Support picture	QCIF, sub-QCIF 10fps
Encoding Options	advanced prediction mode, PB-frame mode, unrestricted vector mode, syntax-based arithmetic coding mode

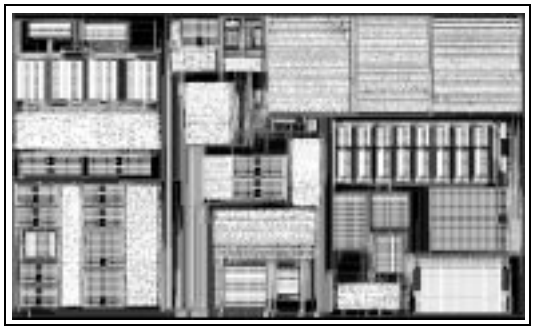


Figure 12: Layout patterns of H.263 codec core (3.3 × 1.9 mm^2).

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 audio-visual communication.

REFERENCES

[1] ITU-T Rec. H. 324: "Terminal for low bit rate multimedia communication", Draft International Standard, Nov. 1995.

[2] ITU-T Rec. H. 263: "Video coding for low bit rate communication", Draft International Standard, May 1996.

[3] ITU-T Rec. H. 261: "Video codec for audiovisual services at $p \times 64$ kbits", International Standard, Mar. 1993.

[4] ISO/IEC 11172-2: "Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s", International Standard, Aug. 1993.

[5] D. Brinthaupt et al.: "A programmable audio/video processor for H.320, H.324, and MPEG", in *IEEE ISSCC Digest of Technical Papers*, pp. 244–245, Feb. 1996.

[6] J. Golston: "Single-chip H.324 videoconferencing", *IEEE Micro*, vol. 16, no. 4, pp. 21–33, Aug. 1996.

[7] M. Harrand et al.: "A single-chip video-phone video encoder/decoder", in *IEEE ISSCC Digest of Technical Papers*, pp. 292–293, Feb. 1996.

[8] K. Okamoto et al.: "A DSP for DCT-based and wavelet-based CODEC's for consumer applications", in *Proc. IEEE Custom Integrated Circuits Conference*, pp. 359–362, May 1996.

[9] Y.-L. Chan and W.-C. Siu: "An efficient block motion vector estimation using adaptive pixel decimation", *Proc. IEEE Int'l Conf. Acoustic, Speech, and Signal Processing*, pp. 2257–2260, May 1995.

[10] M. C. Chen and A. N. Willson Jr.: "A high-accuracy predictive vector quantization motion estimation algorithm for video coding", in *Proc. IEEE Int'l Symp. Circuits and Systems*, pp. 617–620, May 1995.

[11] Y. Kim et al.: "A block matching algorithm with H.26: sub-sampling and its hardware design", in *Proc. IEEE Int'l Symp. Circuits and Systems*, pp. 613–616, May 1995.

[12] N. Hayashi et al.: "A bi-directional motion compensation LSI with a compact motion estimator", *IEICE Trans. Electron.*, vol. E78-C, no. 12, pp. 1682–1690, Dec. 1995.

[13] T. Onoye et al.: "A VLSI-suited motion estimation algorithm based on macroblock cluster matching", *IEICE Tech. Report*, CAS95-45, Jun. 1995 (in Japanese).

[14] T. Onoye et al.: "Single-chip implementation of motion estimator dedicated to MPEG2 MP@HL", *IEICE Trans. Fundamental*, vol. E79-A, no. 8, pp. 1210–1217, Aug. 1996.

[15] G. Fujita et al.: "A new motion estimation core dedicated to H.263 video coding", in *Proc. Int'l Symp. Circuits and Systems*, Jun. 1997, to appear.

[16] ITU-T Rec. H.262 ISO/IEC 13818-2: "Generic coding of moving pictures and associated audio", Draft International Standard, Oct. 1994.

[17] S. Uramoto et al.: "A 100MHz 2-D discrete cosine transform core processor", *IEEE J. Solid-State Circuits*, vol. 27, no. 4, pp. 492–499, Apr. 1992.

[18] M. Matsui et al.: "A 200MHz 13mm^2 2-D DCT macrocell using sense-amplifying pipeline flip-flop scheme", *IEEE J. Solid-State Circuits*, vol. 29, no. 12, pp. 1482–1490, Dec. 1994.

[19] T. Masaki et al.: "VLSI implementation of inverse discrete cosine transformer and motion compensator for MPEG2 HDTV video decoding", *IEEE Trans. Circuits and Systems for Video Technology*, vol. 5, no. 5, pp. 387–395, Oct. 1995.

[20] W. H. Chen et al.: "A fast computational algorithm for the discrete cosine transform", *IEEE Trans. Communications*, vol. COM-25, no. 9, pp. 1004–1009, Sep. 1977.

[21] A. Peled and B. Liu: "A new hardware realization of digital filters", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-22, no. 6, pp. 456–462, Dec. 1974.

[22] E. Tanaka and T. Enomoto: "70mW variable length codec for MPEG2", in *Proc. IEICE General Conference*, C-577, Mar. 1995.