A Two-Chip Real-Time MPEG2 Video Encoder with Wide Range Motion Estimation

Toshio KONDO, Kazuhito SUGURI, Mitsuo IKEDA, Tetsuya ABE, Hiroaki MATSUDA, Tsuneo OKUBO, Kenji OGURA, Yutaka TASHIRO*, Naoki ONO*, Toshihiro MINAMI, Ritsu KUSABA, Takeshi IKENAGA, Nobutaro SHIBATA and Ryota KASAI

NTT LSI Laboratories  *NTT Human Interface Laboratories
Nippon Telegraph and Telephone corporation

Koji OTSU, Fumiaki NAKAGAWA and Yasuhiko SATO
NTT Electronics Technology

3-1, Morinosato Wakamiya, Atsugi-shi, Kanagawa Pref, 243-01, Japan
TEL:+81-468-40-2263  FAX:+81-462-40-2151  E-mail:kond@aecl.ntt.jp

Key Features

• Low-delay SP@ML video encoder

• Wide-range motion estimation using hierarchical telescopic search

• Flexible macroblock-level pipeline architecture based on RISC CPUs

• Three small peripheral memories
### Comparison between SP@ML and MP@ML

<table>
<thead>
<tr>
<th>Items</th>
<th>SP@ML</th>
<th>MP@ML (M =3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delay due to Motion Estimation</td>
<td>&gt; 0.5 Frame**</td>
<td>&gt; 3 Frames</td>
</tr>
<tr>
<td>Frame Memory Capacity</td>
<td>10 Mbits</td>
<td>26 Mbits</td>
</tr>
<tr>
<td>Frame Memory Access Rate*</td>
<td>133 MBytes/s</td>
<td>332 MBytes/s</td>
</tr>
<tr>
<td>Operating Performance Required for Motion Estimation*</td>
<td>3.2 GOPS</td>
<td>8.0 GOPS</td>
</tr>
<tr>
<td>Degradation Compared to MP@ML</td>
<td>SNR at 4 Mbps</td>
<td>—</td>
</tr>
<tr>
<td>Quality for Actual NTSC-Grade Video</td>
<td>Negligible</td>
<td>—</td>
</tr>
</tbody>
</table>

* Hierarchical telescopic search used by our encoder chip set  
** Frame structure

### Comparison of Moving Vector Search Methods

<table>
<thead>
<tr>
<th>Items</th>
<th>Two-Step Exhaustive Search*</th>
<th>Hierarchical Telescopic Search</th>
</tr>
</thead>
<tbody>
<tr>
<td>Required Operating Performance</td>
<td>100 GOPS</td>
<td>3.2 GOPS</td>
</tr>
<tr>
<td>Required Nominal Memory Access Rate</td>
<td>155 MBytes/S</td>
<td>133 MBytes/S</td>
</tr>
<tr>
<td>Signal to Noise Ratio</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flower Garden</td>
<td>28.8 dB</td>
<td>28.6 dB</td>
</tr>
<tr>
<td>Mobile &amp; Calender</td>
<td>26.9 dB</td>
<td>26.7 dB</td>
</tr>
<tr>
<td>Football</td>
<td>35.1 dB</td>
<td>34.8 dB</td>
</tr>
</tbody>
</table>

* First step : full-pel-precision full search  
Second step : half-pel-precision full search
Three-Step Hierarchical Motion Estimation

Previous Field
-3 -2 -1 0

STEP 1
Telescopic Search by 8 x 4 PEs

STEP 2
Full-pel Search by 3 x 3 PEs

STEP 3
Half-pel Search by 8 PEs

Telescopic search -

Template Macroblock

(1)

(4) (3) (2)

(5) (6)

Reduction of Frame Memory Access Rate by Full Buffering

No Buffering

Previous Search Area

Current Search Area

MB MB

Current Search Area = 3 x MB

Buffering of Total Possible Search Area

Previous Possible Search Area

Current Possible Search Area

New Area

MB MB

New Area = 5 x MB
Function Partitioning in MPEG2 Encoding Procedure

Reordering Field Buffer (4Mbit VRAM)

Video Input [4:2:2]

Noise Filter

MB Type & Activity Detection

Rate & Buffer Control

DCT

Q

VLC

Output Buffer (2Mbit FIFO DRAM)

Bit Stream

ENC - M

ENC - C

Frame Memory (16Mbit SDRAM or 4Mbit SDRAM x2)

Flexible Pipeline Architecture Based on RISC CPUs

Video Input

VRAM I/F

VRAM

Motion Estimation & Motion Compensation

DCT I/F

DCT

DCT/Q

VLC

FIFO DRAM

Host I/F

Host CPU
Diagram of Motion Compensation Chip (ENC-M)

Program Memory  Data Memory  RISC CPU-M  BUS I/F


VRAM I/F  SDRAM I/F

VRAM  SDRAM

d to ENC-C
d to ENC-C

Layout of ENC-M Chip

- 0.5-μm triple-metal CMOS
- 2.0 million transistors
  Logic : 150k G
  Memory:
  Program 24bx3.5kw
  Data 16bx1kw
  Others 91kb
- 16.5 x 16.5 mm² die size
- 81-MHz clock
- 3.5 Watts @ 3.3 V
- 340-lead CQFP
Layout of ENC-C Chip

- 0.5-μm triple-metal CMOS
- 1.3 million transistors
  Logic: 75k G
  Memory:
  Program 24bx5kw
  Data 10bx1kw
  Others 20Kb
- 14.0 x 14.0 mm² die size
- 81-MHz clock
- 2.5 Watts @ 3.3 V
- 204-lead CQFP

Video Encoder Specifications

<table>
<thead>
<tr>
<th>Standard</th>
<th>MPEG2</th>
<th>SP@ML</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frame Size</td>
<td>720 x 480 / 720 x 576 / 640 x 480 (variable)</td>
<td></td>
</tr>
<tr>
<td>Video Input Format</td>
<td>4 : 2 : 2</td>
<td></td>
</tr>
<tr>
<td>Frame Rate</td>
<td>Up to 30 frames/s (Up to 40,500 macroblocks/s)</td>
<td></td>
</tr>
<tr>
<td>Bit Rate</td>
<td>Up to 15 Mbps</td>
<td></td>
</tr>
<tr>
<td>Picture Structure</td>
<td>Frame Structure</td>
<td></td>
</tr>
<tr>
<td>Motion Estimation Search Range</td>
<td>Frame: -32/31.5 h x -32/31.5 v</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Field: -48/47.5 h x -24/23.5 v</td>
<td></td>
</tr>
<tr>
<td>Encoder Delay</td>
<td>85 ms</td>
<td></td>
</tr>
</tbody>
</table>