A 0.9-V, 150-MHz, 10-mW, 4 mm², 2-D Discrete Cosine Transform Core Processor with Variable Threshold-Voltage (VT) Scheme

Tadahiro Kuroda, *Member, IEEE*, Tetsuya Fujita, Shinji Mita, Tetsu Nagamatsu, Shinichi Yoshioka, Kojiro Suzuki, Fumihiko Sano, Masayuki Norishima, Masayuki Murota, Makoto Kako, Masaaki Kinugawa, *Member, IEEE*, Masakazu Kakumu, *Member, IEEE*, and Takayasu Sakurai, *Member, IEEE*

Abstract— A 4 mm², two-dimensional (2-D) 8 × 8 discrete cosine transform (DCT) core processor for HDTV-resolution video compression/decompression in a 0.3-μm CMOS triple-well, double-metal technology operates at 150 MHz from a 0.9-V power supply and consumes 10 mW, only 2% power dissipation of a previous 3.3-V design. Circuit techniques for dynamically varying threshold voltage (VT scheme) are introduced to reduce active power dissipation with negligible overhead in speed, standby power dissipation, and chip area. A way to explore $V_{DD} - V_{th}$ design space is also studied.

I. INTRODUCTION

LOWERING both the supply voltage V_{DD} and threshold voltage V_{th} enables high-speed, low-power operation [1], [2]. This approach, however, raises two problems [3], [4], 1) degradation of worst-case speed due to V_{th} fluctuation in low V_{DD}, and 2) increase in standby power dissipation in low V_{th}. To solve these problems, several schemes are proposed. A self-adjusting threshold voltage (SAT) scheme [5] reduces V_{th} fluctuation in an active mode by adjusting substrate bias with a feedback control circuit. A standby power reduction (SPR) scheme [6] raises V_{th} in a standby mode by switching substrate bias between the power supply and an external additional supply higher than V_{DD} or lower than GND. A multi threshold voltage CMOS (MT-CMOS) scheme [7] employ low V_{th} for fast circuit operation and high V_{th} for providing and cutting internal supply voltage. The SAT and the SPR are both based upon the same idea that V_{th} is controlled dynamically through substrate bias. However, the two schemes cannot be combined because the SPR requires the external supply for the substrate bias while the SAT generates the substrate bias internally. The MT-CMOS does not solve the first problem. It requires very large transistors for the internal power supply control to impose area and yield penalties, otherwise degrading circuit speed. Furthermore, it cannot be applied to memory

elements without circuit tricks which add another area and speed penalties.

This paper presents a variable threshold voltage scheme (VT scheme) which can solve these two problems uniformly in a unified way by controlling substrate bias with substrate bias feedback control circuits. Unlike the conventional approaches, it requires no external power supply for the substrate bias, leaves no restriction in use, imposes practically no penalty in speed and chip area, and can be applied to both logic gates and memory elements. The VT scheme is employed in a two-dimensional (2-D) 8 × 8 discrete cosine transform (DCT) core processor for portable HDTV-resolution video compression/decompression. This DCT in a 0.3-μm CMOS technology operates at 150 MHz from a 0.9-V power supply and consumes 10 mW, only 2% power dissipation of a previous 3.3-V design [8].

In Section II, low V_{DD}, low V_{th} design space is explored to investigate V_{th} target. In Section III, the VT scheme is presented, followed by descriptions of circuit implementations in Section IV. Section V details the design of the DCT. Experimental results appear in Section VI. Section VII is dedicated for conclusions.

II. EXPLORING LOW-V_{DD} LOW-V_{th} DESIGN SPACE

CMOS power dissipation is given by

$$P = \tfrac{1}{2} \cdot p_t \cdot f_{CLK} \cdot C_L \cdot V_{DD}^2 + I_0 \cdot 10^{-(V_{th}/S)} \cdot V_{DD} \quad (1)$$

where p_t is the switching probability, f_{CLK} is the clock frequency, C_L is the load capacitance, S is the subthreshold swing, and I_0 is a constant which is proportional to total transistor width in a chip. The first term represents dynamic power dissipation due to charging and discharging of the load capacitance, and the second term is leakage current dissipation due to subthreshold conduction. Since the dominant term in a typical CMOS design is the dynamic power dissipation, lowering V_{DD} is effective to low-power design.

Gate propagation delay, on the other hand, is approximately given in [9] by

$$t_{pd} = \frac{k \cdot C_L \cdot V_{DD}}{(V_{DD} - V_{th})^\alpha} \quad (2)$$

where α is typically 1.3 and k is a constant. Lowering only V_{DD} leads to slower circuit speed, and therefore, both of V_{DD} and V_{th} should be lowered for high-speed, low-power design. When V_{DD} and V_{th} are lowered to V'_{DD} and V'_{th}, and the circuit speed becomes λ times, their relation is given from (2) by

$$\frac{V'_{DD} - V'_{th}}{V_{DD} - V_{th}} = \left(\lambda \cdot \frac{V'_{DD}}{V_{DD}} \right)^{1/\alpha}. \tag{3}$$

For example, suppose $V_{DD} = 3.3$ V and $V_{th} = 0.6$ V. Under a constant speed condition ($\lambda = 1$), one solution is $V'_{DD} = 2.1$ V and $V'_{th} = 0.2$ V. In this case, the dynamic power dissipation is reduced to 41%. If circuit speed can be reduced to 60% ($\lambda = 0.6$), the dynamic power dissipation can be reduced to 7% at $V'_{DD} = 0.9$ V and $V'_{th} = 0.2$ V.

For more precise estimation, process fluctuation should be taken into account. V_{th} fluctuates typically by ± 0.1 V, which causes t_{pd} variation. From (2), the variation in t_{pd}, K_{VT}, is given by

$$K_{VT} = \frac{\Delta t_{pd}}{t_{pd}}$$
$$= \frac{\alpha \cdot \Delta V_{th}}{V_{DD} - V_{th}}. \tag{4}$$

In order to assure high yield in production, margin should be incorporated into design so as to satisfy speed specification even with fluctuations in process. Smaller K_{VT} leads to smaller design margin, and therefore, is preferable from area-saving and low-power design point of view. In lowering both V_{DD} and V_{th}, K_{VT} should be kept at least from increasing. From (3) and (4), calculating the condition to keep K_{VT} constant yields

$$\frac{\Delta V'_{th}}{\Delta V_{th}} = \left(\lambda \cdot \frac{V'_{DD}}{V_{DD}} \right)^{1/\alpha}. \tag{5}$$

In the former examples, under the constant speed condition ($\lambda = 1$) V_{th} fluctuation should be reduced to $\Delta V'_{th}/\Delta V_{th} = 0.71$, and in the 60%-speed condition ($\lambda = 0.6$) it should be reduced to $\Delta V'_{th}/\Delta V_{th} = 0.25$.

But in reality, it is not expected that as V_{th} is lowered, ΔV_{th} is reduced as much. If impurity density in the channel region is simply reduced to lower V_{th} of a surface channel device such as nMOS with n$^+$ polysilicon gates, the short-channel effect degrades to increase ΔV_{th}, reflecting variation of polysilicon gates in size. In a buried-channel device such as pMOS with n$^+$ polysilicon gates, on the other hand, counter doping should be added to lower V_{th}, resulting in higher impurity density and larger ΔV_{th}. It is not that simple to discuss ΔV_{th}, but generally speaking, device researchers expect ΔV_{th} could be increased in low V_{th} and would not be decreased very easily. This is one issue in low-V_{DD}, low-V_{th} CMOS circuit design.

Another issue is the rapid increase in subthreshold leakage in low V_{th} as seen from (1). In portable applications it is clear that large standby leakage becomes a problem. Not only in portable applications but also in desktop applications, the rapid increase in subthreshold current determines the lower limit of V_{th}, and therefore, it is also important.

In order to study these two issues and explore low-V_{DD}, low-V_{th} design space, (1) and (2) are numerically solved with the parameters for this DCT design in a 0.3-μm CMOS technology at junction temperature of 90°C. Contour lines in terms of speed (i.e., maximum operating frequency) and power are drawn on the $V_{DD} - V_{th}$ plane in Fig. 1. In a typical 3.3-V design, V_{DD} is at 3.3 $\pm$ 10% V and V_{th} is set to 0.6 $\pm$ 0.1 V. The design space is represented by a box in Fig. 1. The maximum operating frequency, f, becomes the slowest, 250 MHz, at $V_{DD} = 3.0$ V and $V_{th} = 0.7$ V. The circuit speed is therefore normalized ($\lambda = 1$) at the upper-left corner of the design-space box. The power dissipation, on the other hand, becomes the largest, 160 mW, at $V_{DD} = 3.6$ V and $V_{th} = 0.5$ V. The power dissipation is therefore normalized ($\xi = 1$) at the lower-right corner of the design-space box. For designing a 150 MHz DCT, which is 60% speed of 250 MHz, the upper-left corner of the design-space box should be placed on the speed contour line with $\lambda = 0.6$. It is found from the lower-right corner of the design-space box that the power dissipation can be reduced to 25% ($\xi = 0.25$), that is 40 mW, by lowering V_{DD} to 1.9 $\pm$ 10% V. It can further be reduced to 6% ($\xi = 0.06$), that is 10 mW, by lowering both V_{DD} to 1.0 $\pm$ 10% V and V_{th} to 0.27 $\pm$ 0.02 V. This supply voltage can be supplied from a single battery source. Reducing ΔV_{th} from ± 0.1 V to ± 0.02 V also meets the requirement for keeping K_{VT} constant in (5). As shown in Fig. 1, power dissipation due to subthreshold leakage becomes about 1% of the total power dissipation.

To summarize, V_{DD} should be at 1.0 $\pm$ 10% V, and V_{th} should be controlled at 0.27 $\pm$ 0.02 V in the active mode and higher than 0.5 V in the standby mode.

III. VARIABLE THRESHOLD-VOLTAGE (VT) SCHEME

The VT scheme is conceptually illustrated in Fig. 2. Threshold voltage of a transistor is variable through substrate bias control with a Variable Threshold-voltage circuit (VT circuit). In the active mode, the VT circuit controls the substrate bias, V_{BB}, so as to compensate the V_{th} fluctuation. Even though device V_{th} has 0.1-V fluctuation around 0.15 V, V_{th} is compensated and set at 0.27 $\pm$ 0.02 V in the active mode. In the standby mode, the VT circuit applies deeper substrate bias to increase V_{th} to higher than 0.5 V and cut off leakage. Typically, V_{BB} of -0.5 V is applied in the active mode and -3.3 V in the standby mode.

Fig. 3 depicts the VT scheme block diagram. The VT scheme consists of four leakage current monitors (LCM's), the self-substrate bias circuit (SSB), and a substrate charge injector (SCI). The SSB draws current from the substrate to lower V_{BB}. The SCI, on the other hand, injects current into the substrate to raise V_{BB}. The SSB and the SCI are controlled by monitoring where V_{BB} sits in four ranges. Their criteria are specified in the four LCM's; $V_{active(+)} = -0.3$ V, $V_{active} = -0.5$ V, $V_{active(-)} = -0.7$ V, and $V_{standby} = -3.3$ V. The substrate bias is monitored by transistor leakage current, because the leakage current reflects V_{BB} very sensitively.

Fig. 4 illustrates the substrate bias control. After a power-on, V_{BB} is higher than $V_{active(+)}$, and the SSB begins to draw
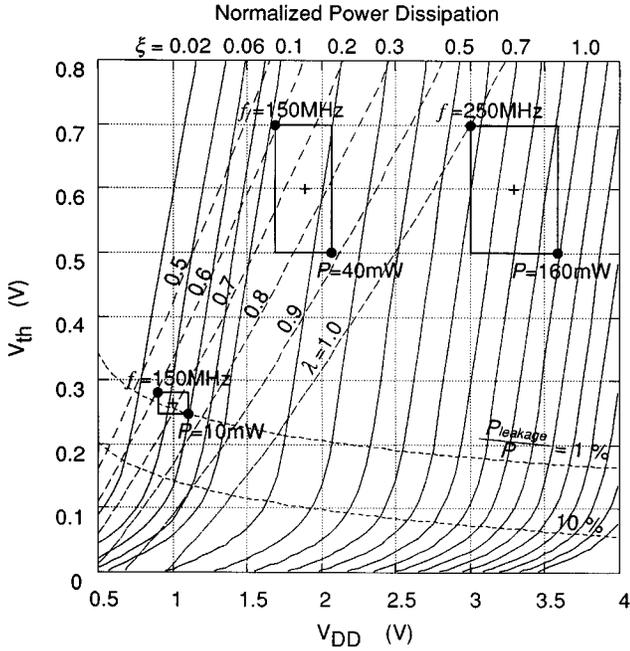
Fig. 1. Exploring low-V_{DD}, low-V_{th} design space. Contour lines in terms of speed (broken lines) and power (solid lines) are drawn.

Fig. 3. VT block diagram.

Fig. 4. Substrate-bias control in VT.

Fig. 2. Variable threshold-voltage (VT) scheme.

100 μA from the substrate to lower V_{BB} using a 50 MHz ring oscillator. This current is large enough for V_{BB} to settle down within 10 μs after a power-on. When V_{BB} goes lower than $V_{\mathrm{active}(+)}$, the pump driving frequency drops to 5 MHz and the SSB draws 10 μA to control V_{BB} more precisely. The SSB stops when V_{BB} drops below V_{active}. V_{BB}, however, rises gradually due to device leakage current through MOS transistors and junctions, and reaches V_{active} to activate the SSB again. In this way, V_{BB} is controlled at V_{active} by the on–off control of the SSB. When V_{BB} goes deeper than $V_{\mathrm{active}(-)}$, the SCI turns on to inject 30 mA into the substrate. Therefore, even if V_{BB} jumps beyond $V_{\mathrm{active}(+)}$ or $V_{\mathrm{active}(-)}$ due to a power line bump for example, V_{BB} is quickly recovered to V_{active} by the SSB and the SCI. When "SLEEP" signal is asserted ("1") to go to the standby mode, the SCI is disabled and the SSB is activated again and 100 μA current is

drawn from the substrate until V_{BB} reaches V_{standby}. V_{BB} is controlled at V_{standby} in the same way by the on–off control of the SSB. When "SLEEP" signal becomes "0" to go back to the active mode, the SSB is disabled and the SCI is activated. The SCI injects 30 mA current into the substrate until V_{BB} reaches $V_{\mathrm{active}(-)}$. V_{BB} is finally set at V_{active}. In this way, the SSB is mainly used for a transition from the active mode to the standby mode, while the SCI is used for a transition from the standby to the active mode. An active to standby mode transition takes about 100 μs, while a standby to active mode transition is completed in 0.1 μs. This "slow falling asleep but fast awakening" feature is acceptable for most of the applications.

The SSB operates intermittently to compensate for the voltage fluctuation in the substrate due to the substrate current in the active and the standby modes. It therefore consumes several microamperes in the active mode and less than one nanoampere in the standby mode, both much lower than the chip power dissipation. Energy required to charge and discharge the substrate for switching between the active and the standby modes is less than 10 nJ. Even when the mode is switched 1000 times in a second, the power dissipation becomes only 10 μW. The leakage current monitor should be designed to dissipate less than 1 nA because it always works even in the standby mode. The low-power circuit design technique is described in the next section.

IV. CIRCUIT IMPLEMENTATIONS

A. Leakage Current Monitor (LCM)

The substrate bias is generated by the SSB which is controlled by the leakage current monitor (LCM). The LCM is therefore a key to the accurate control in the VT scheme. Fig. 5
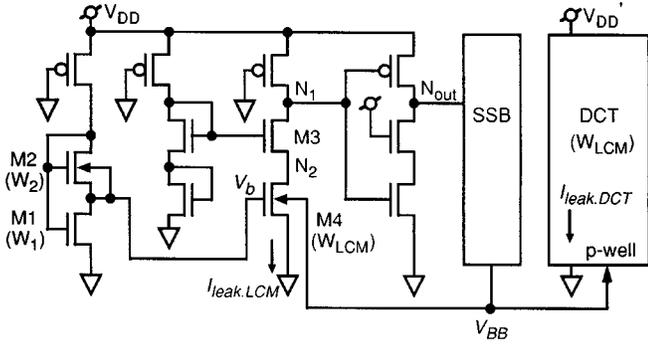
Fig. 5. Leakage current monitor (LCM).

depicts a circuit schematic of the proposed LCM. The circuit works with 3.3-V V_{DD} which is usually available on a chip for standard interfaces with other chips. The LCM monitors leakage current of the DCT, $I_{\text{leak.DCT}}$, with a transistor M4 that shares the same substrate with the DCT. The gate of M4 is biased to V_b to amplify the monitored leakage current, $I_{\text{leak.LCM}}$. If $I_{\text{leak.LCM}}$ is larger than a target reflecting shallower V_{BB} and lower V_{th}, the node N_1 goes "*Low*" and the output node N_{out} goes "*High*" to activate the SSB. As a result, V_{BB} goes deeper and V_{th} becomes higher, and consequently, $I_{\text{leak.LCM}}$ and $I_{\text{leak.DCT}}$ become smaller. When $I_{\text{leak.LCM}}$ becomes smaller than the target, the SSB stops. Then $I_{\text{leak.LCM}}$ and $I_{\text{leak.DCT}}$ increase as V_{BB} gradually rises due to device leakage current through MOS transistors and junctions, and finally reaches the target to activate the SSB again. In this way, $I_{\text{leak.DCT}}$ is set to a target by the on–off control of the SSB with the LCM.

In order to make this feedback control accurately, the current ratio of $I_{\text{leak.LCM}}$ to $I_{\text{leak.DCT}}$, or the current magnification factor of the LCM, X_{LCM}, should be constant. When an MOS transistor is in subthreshold, its drain current is expressed as

$$I_{DS} = \frac{I_0}{W_0} \cdot W \cdot 10^{(V_{GS} - V_T)/S} \qquad (6)$$

where S is the subthreshold swing, V_T is the threshold voltage, I_0/W_0 is the current density to define V_T, and W is the channel width. By applying (6), X_{LCM} is given by

$$X_{\text{LCM}} = \frac{I_{\text{leak.LCM}}}{I_{\text{leak.DCT}}}$$
$$= \frac{W_{\text{LCM}}}{W_{\text{DCT}}} \cdot 10^{V_b/S} \qquad (7)$$

where W_{DCT} is the total channel width of the DCT and W_{LCM} is the channel width of M4. Since two transistors M1 and M2 in a bias generator are designed to operate in subthreshold region, the output voltage of the bias generator V_b is also given from (6) by

$$V_b = S \cdot \log \frac{W_2}{W_1} \qquad (8)$$

where W_1 and W_2 is the channel width of M1 and M2, respectively. X_{LCM} is therefore expressed as

$$X_{\text{LCM}} = \frac{W_2}{W_1} \cdot \frac{W_{\text{LCM}}}{W_{\text{DCT}}}. \qquad (9)$$
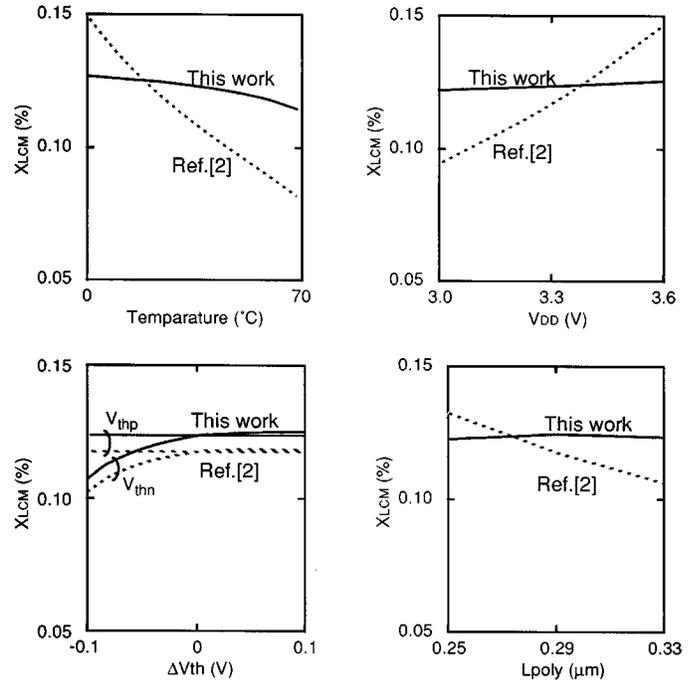
Fig. 6. Current magnification factor of the LCM, X_{LCM}, dependence on circuit condition changes and process deviations simulated by SPICE.

This implies that X_{LCM} is determined only by the transistor size ratio and independent of the power supply voltage, temperature, and process fluctuation. In the conventional circuit [5], on the other hand, where V_b is generated by dividing the V_{DD}-GND voltage with high impedance resistors, V_b becomes a function of V_{DD}, and therefore, X_{LCM} becomes a function of V_{DD} and S, where S is a function of temperature. Fig. 6 shows SPICE simulation results of X_{LCM} dependence on circuit condition changes and process fluctuation. X_{LCM} exhibits small dependence on ΔV_{thn} and temperature. This is because M4 is not in deep subthreshold region. The variation of X_{LCM}, however, is within 15%, which results in less than 1% error in V_{th} controllability. This is negligible compared to 20% error in the conventional implementation.

The four criteria used in the substrate-bias control, corresponding to $V_{\text{active}(+)}$, V_{active}, $V_{\text{active}(-)}$, and V_{standby} can be set in the four LCM's by adjusting the transistor size W_1, W_2, and W_{LCM} in the bias circuit. For the active mode, with $W_1 = 10\,\mu\text{m}$, $W_2 = 100\,\mu\text{m}$, and $W_{\text{LCM}} = 100\,\mu\text{m}$, the magnification factor X_{LCM} of 0.001 is obtained when $W_{\text{DCT}} = 1$ m. $I_{\text{leak.DCT}}$ of 0.1 mA can be monitored as $I_{\text{leak.LCM}}$ of 0.1 μA in the active mode. For the standby mode, with $W_1 = 10\,\mu\text{m}$, $W_2 = 1000\,\mu\text{m}$, and $W_{\text{LCM}} = 1000\,\mu\text{m}$, X_{LCM} becomes 0.1. Therefore, $I_{\text{leak.DCT}}$ of 10 nA can be monitored as $I_{\text{leak.LCM}}$ of 1 nA in the standby mode. The overhead in power by the monitor circuit is about 0.1 and 10% of the total power dissipation in the active and the standby mode, respectively.

The parasitic capacitance at the node N_2 is large because M4 is large. This may degrade response speed of the circuit. The transistor M3, however, isolates the N_1 node from the N_2 node and keeps the signal swing on N_2 very small. This reduces the response delay and improves dynamic V_{th} controllability.

Fig. 7. Pump circuit in SSB.

Fig. 8. SCI and its waveforms simulated by SPICE.
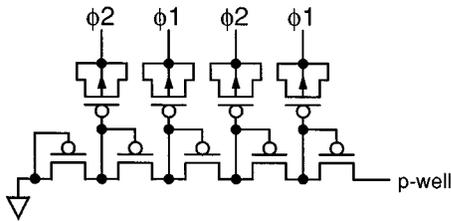
Compared with the conventional LCM where V_b is generated by dividing the V_{DD}-GND voltage with high impedance resistors, the V_{th} controllability including the static and dynamic effects is improved from ±0.05 V to less than ±0.01 V, response delay is shortened from 0.6 to 0.1 μs, and Si area is reduced from 33 250 to 670 μm^2. This layout area reduction is brought by the elimination of the high impedance resistors by polysilicon.

B. Self-Substrate Bias Circuit (SSB)

Fig. 7 depicts a schematic diagram of a pump circuit in the SSB. PMOS transistors of the diode configuration are connected in series whose intermediate nodes are driven by two signals, $\Phi1$ and $\Phi2$, in 180° phase shift. Every other transistor, therefore, sends current alternately from p-well to GND, resulting in lower p-well bias than GND. The SSB can pump as low as -4.5 V. SSB circuits are widely used in DRAM's and E^{2}PROM's, but two orders of magnitude smaller circuit can be used in the VT scheme. The driving current of the SSB is 100 μA, while it is usually several milliamperes in DRAM's. This is because substrate current generation due to the impact ionization is a strong function of the supply voltage. Substrate current in a 0.9-V DCT is considerably smaller than that in a 3.3-V design. Substrate current introduced from I/O pads does not affect the DCT macro because it is separated from peripheral circuits by a triple-well structure. Eventually, no substrate current is generated in the standby mode. From these reasons, the pumping current in the SSB can be as small as several percent of that in DRAM's. Silicon area is also reduced considerably. Another concern about the SSB is an initialization time after a power-on. Even in a 10 mm square chip, V_{BB} settles down within 200 μs after a power-on, which is acceptable in real use.

C. Substrate Charge Injector (SCI)

In the VT scheme, care should be taken so that no transistor sees high-voltage stress of gate oxide and junctions. Transistors are optimized for use at 3.3 V. The gate oxide thickness is 8 nm. The maximum voltage that assures sufficient reliability of the gate oxide is $V_{DD}+20\%$, or 4 V. The SCI in Fig. 8 receives a control signal that swings between V_{DD} and GND at node N_1 to drive substrate from V_{standby} to V_{active}. In the standby-to-active transition, $V_{DD} + |V_{\text{standby}}|$ that is about 6.6 V at maximum can be applied between N_1 and N_2. However, as shown in SPICE simulated waveforms in Fig. 8, $|V_{GS}|$ and $|V_{GD}|$ of M1 and M2 never exceeds the larger of V_{DD} and $|V_{\text{standby}}|$. All other transistors in the VT circuit and
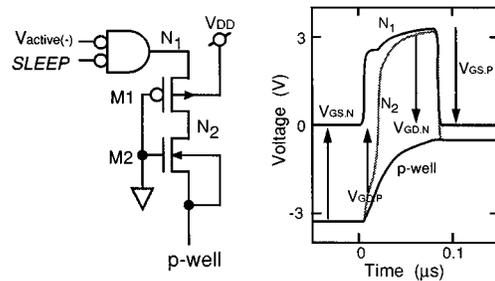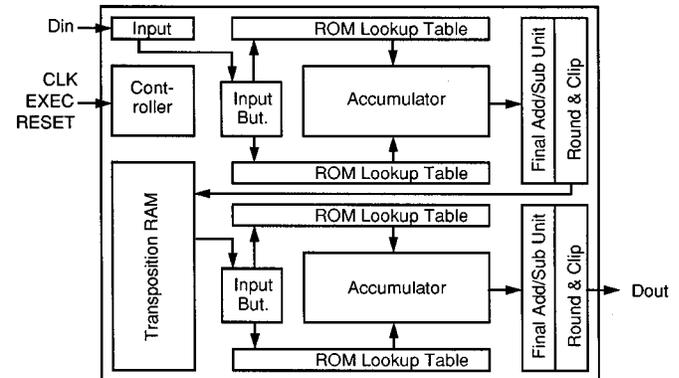
Fig. 9. DCT block diagram.

the DCT macro receive $(V_{DD} - V_{th})$ on their gate oxide when the channel is formed in the depletion and the inversion mode, and less than $|V_{\text{standby}}|$ in the accumulation mode. These considerations lead to a general guideline that V_{standby} should be limited to $-(V_{DD} + 20\%)$. V_{standby} of $-(V_{DD} + 20\%)$, however, can shift V_{th} big enough to reduce the leakage current in the standby mode. The body effect coefficient, γ, can be adjusted independently to V_{th} by controlling the doping concentration density in the channel-substrate depletion layer.

V. DCT DESIGN

A. Circuit Design

This DCT core processor executes 2-D 8 × 8 DCT and inverse DCT. A block diagram is illustrated in Fig. 9. The DCT is composed of two one-dimensional (1-D) DCT and inverse DCT processing units and a transposition RAM. Rounding circuits and clipping circuits which prevent overflow and underflow are also implemented in the cell. The DCT has a concurrent architecture based on distributed arithmetic and a fast DCT algorithm, which enables high throughput DCT processing of one pixel per clock. It also has fully pipelined structure. The 64 input data sampled in every clock cycles are outputted after 112 clock cycle latency.

Various memories which use the same low V_{th} transistors as logic gates are employed in the DCT. Table lookup ROM's (16 b × 32 words × 16 banks) employ contact programming and an inverter-type sense-amplifier. Single-port SRAM's (16 b × 64 words × 2 banks) and dual-port SRAM's (16 b × 8 words × 2 banks) employ a six-transistor cell and a latch sense-amplifier. They all exhibit wide operational margin in
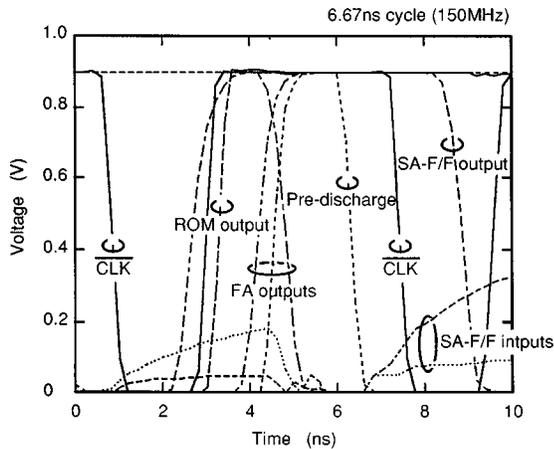
Fig. 10. Simulated waveforms of MAC datapath.

low V_{DD} and low V_{th} and almost behave like logic gates in terms of circuit speed dependence on V_{DD} and V_{th}. No special care is necessary such as word-line boosted-up or a special sense-amplifier.

Small-swing differential pass-transistor logic (SAPL) with sense-amplifying pipeline flip-flop (SA–F/F) [8] is employed for high-speed operation in a 20-b carry skip adder in an accumulator. The SAPL operates and maintains its speed advantage at 0.9 V because the SA–F/F uses a current-mode latch sense-amplifier. As shown in SPICE simulation in Fig. 10, a multiplication and accumulation (MAC) datapath runs at 150 MHz under 0.9 V with no modifications from the 3.3-V design [8].

B. Layout Design

In the conventional CMOS, substrate-contacts are connected to power lines locally, while in the VT scheme they should be interconnected globally for biasing the substrate. This may impose area penalty for separating many substrate contacts, or performance degradation due to substrate noise with few substrate contacts. It is considered, however, that not many substrate contacts are needed in 0.9-V design compared to 3.3-V design because the substrate current generated by the impact ionization becomes several orders of magnitude smaller in 0.9 V. As for substrate noise induced by drain-substrate capacitive coupling, lowering supply voltage is favorable because signal swing as a noise source becomes smaller. It should be effective to add source diffusions because it helps to stabilize V_{BB} by junction capacitance between source diffusions and substrate.

Layout design of the DCT is made in the conventional CMOS fashion, and then it is automatically modified for the VT scheme as illustrated in Fig. 11 by a script of a layout editor. First, have the DCT macro wrapped by deep n-well. Second, generate p-well by inverting n-well data in the deep n-well. Third, replace all the substrate-contacts by source diffusions as long as design rules accept, otherwise remove them. Lastly, place substrate-contacts at the periphery of the deep n-well and the p-well. The p-well becomes one big island and can be connected at periphery. The n-well, on the other hand, becomes many pieces of separated islands.

However, they sit in one deep n-well and can be connected at the periphery, too. Since substrate contacts are only placed at periphery of the 2 mm-square macro, large parasitic substrate resistance is included. Performance degradation or latch up effect due to substrate noise should be examined. Experimental results are presented in the next section. The area penalty, on the other hand, becomes less than 0.1%. This can be done in a p-well or an n-well technology, too, but triple-well structure prevents I/O noise from affecting the DCT macro. The increase in cost and turnaround time by introducing triple-well process is less than 5%. The necessity of the triple-well structure should be examined in the future.

VI. EXPERIMENTAL RESULTS

The DCT core processor is fabricated in a 0.3-μm CMOS, triple-well, double-metal technology. Parameters of the technology and the features of the DCT macro are summarized in Table I. It operates with 0.9-V power supply which can be supplied from a single battery source. Power dissipation at 150 MHz operation is 10 mW. The leakage current in the active mode is 0.1 mA, about 1% of the total power current. The standby leakage current is less than 10 nA, four orders of magnitude smaller than the active leakage current. A chip micrograph appears in Fig. 12(a). The core size is 2 mm square. A magnified picture of the VT control circuit appears in Fig. 12(b). It occupies 0.37 mm $\times$ 0.52 mm, less than 5% of the macro size. If additional circuits for testability are removed and the layout is optimized, the layout size is estimated to be 0.3 mm $\times$ 0.3 mm. The VT circuit is symmetric for p-well and n-well control. LCM(N), however, occupies more area than LCM(P) because nMOS transistor loads in LCM(N) need longer gate length than pMOS transistor loads in LCM(P) for monitoring the same $I_{\text{leak,LCM}}$.

Fig. 13(a)–(c) shows measured p-well voltage waveforms. Due to large parasitic capacitance in a probe card, the transition takes longer time than SPICE simulation. Just after the power-on, the VT circuits are not activated yet because the power supply is not high enough. As shown in Fig. 13(a), p-well is biased forward by 0.2 V due to capacitance coupling between p-well and power lines. Then the VT circuits are activated and p-well is to be biased at −0.5 V. It tales about 8 μs to be ready for the active mode after the power-on. The active-to-standby mode transition takes about 120 μs as shown in Fig. 13(b), while the standby-to-active mode transition is completed within 0.2 μs as presented in Fig. 13(c).

Compared to the DCT in [8], power dissipation at 150 MHz operation is reduced from 500 mW to 10 mW, that is only 2%. Most of the power reduction, however, is brought by capacitance reduction and voltage reduction by technology scaling. Technology scaling from 0.8 to 0.3 μm reduces power dissipation from 500 to 100 mW at 3.3 V and 150 MHz operation. Without the VT scheme, V_{DD} and V_{th} cannot be lowered under 1.7 and 0.5 V, respectively, and the active power dissipation is to be 40 mW. It is therefore fair to claim that the VT scheme contributes to reduce the active power dissipation from 40 to 10 mW.

The DCT operates at supply voltages from 0.9 to above 3 V. No performance degradation nor latchup effect is observed
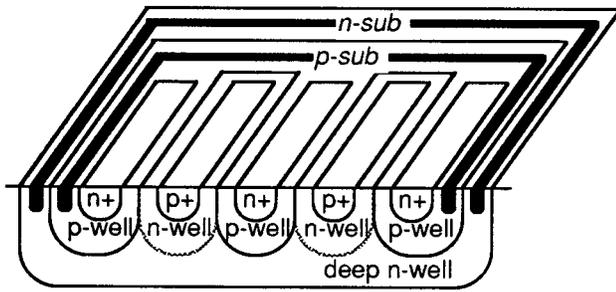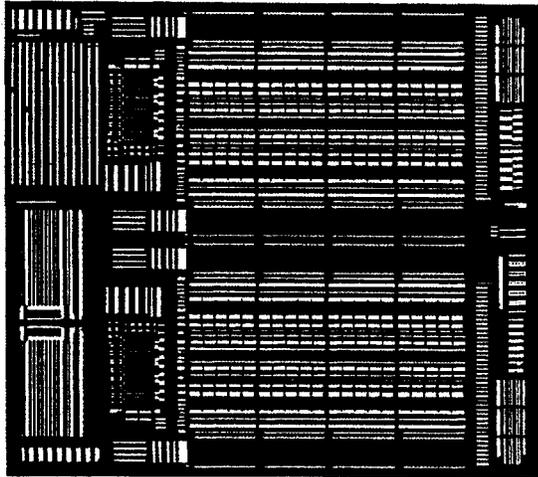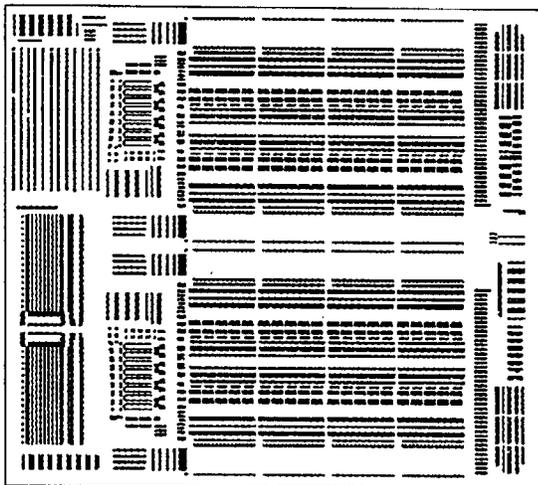
Fig. 11. DCT layout modification for the VT scheme: (a) device cross-section, (b) p-well (one island), and (c) n-well (pieces of islands) in deep n-well.

TABLE I
FEATURES

Technology	0.3 μm CMOS, triple-well, double-metal, $T_{ox} = 8$ nm, $V_{th} = 0.15$ V $\pm$ 0.1 V
Power supply voltage	1.0 V $\pm$ 0.1 V
Power dissipation	10 mW @ 150 MHz
Standby current	<10 nA @ 70°C
Transistor count	120K Tr
Area	2.0×2.0 mm^2
Function	8×8 DCT and inverse DCT
Data format	9-b signed (pixel), 12-b signed (DCT)
Latency	112 clocks
Throughput	64 clocks/block
Accuracy	CCITT H.261 compatible
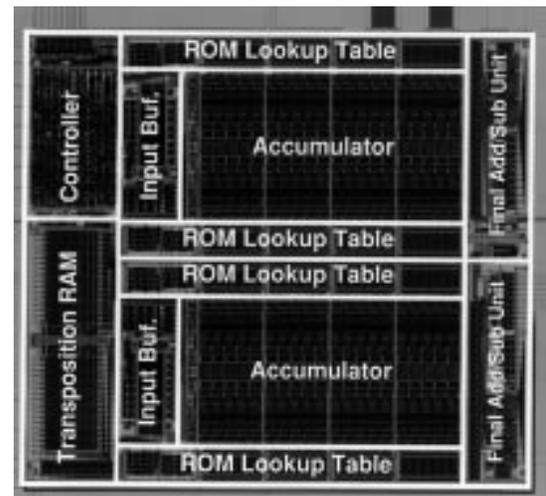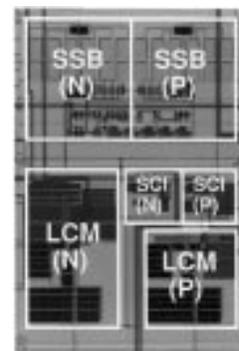
Fig. 12. Chip micrograph: (a) DCT macro and (b) VT circuits.

even when 100 kΩ resistance is added between the substrate and the output of the SSB.

VII. CONCLUSIONS

A 4 mm^2 2-D DCT core processor for portable multimedia equipment with HDTV-resolution video compression and decompression has been developed in a 0.3-μm CMOS, triple-well, double-metal technology. It operates at 150 MHz from a 0.9 V power supply and dissipates 10 mW, which is only 2% of the previous 3.3 V design. Circuit design techniques for dynamically varying threshold voltage (VT scheme) are introduced to reduce active power dissipation with negligible overhead in speed, standby power dissipation, and chip area. The active-to-standby mode transition takes 120 μs, while the standby-to-active mode transition is completed within 0.2 μs. The VT scheme can be applied to both logic gates and memory elements. Generation of the low-voltage V'_{DD} on chip is a future research work.
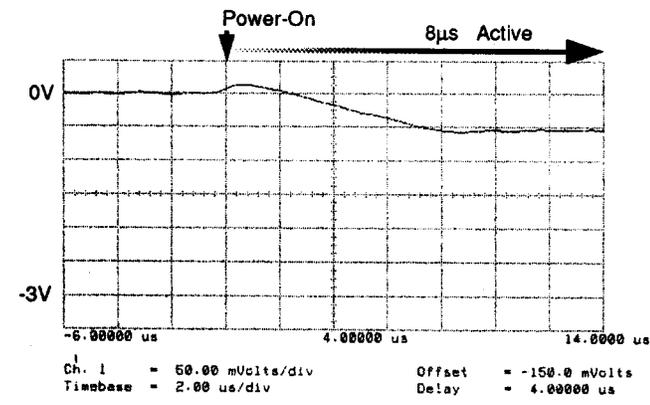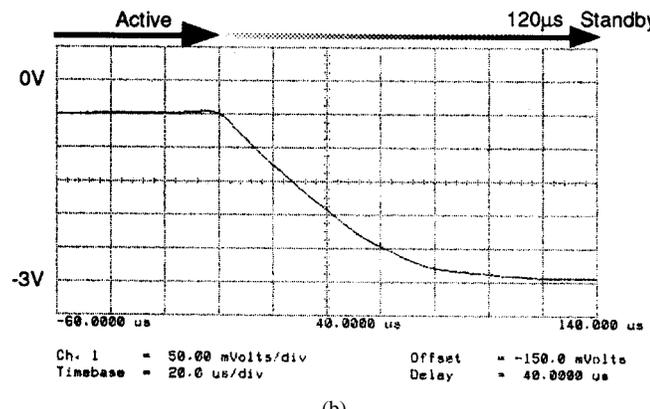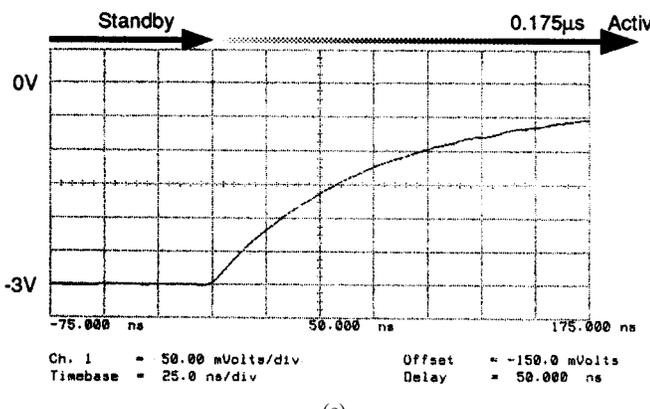
Fig. 13. Measured p-well V_{BB}: (a) after power-on, (b) active-to-standby, and (c) standby-to-active.

REFERENCES

[1] T. Kuroda and T. Sakurai, "Overview of low-power ULSI circuit techniques," *IEICE Trans. Electron.,* vol. E78-C, no. 4, pp. 334–344, Apr. 1995.
[2] J. B. Burr and T. Shott, "A 200 mV self-testing encoder/decoder using Stanford ultra-low-power CMOS," in *ISSCC Dig. Tech. Papers,* Feb. 1994, pp. 84–85.
[3] T. Kuroda and T. Sakurai, "Threshold-voltage control schemes through substrate-bias for low-power high-speed CMOS LSI design," in *J. VLSI Signal Processing.* Norwell, MA: Kluwer, Special Issue on *Technologies for Wireless Computing,* to be published.
[4] S.-W. Sun and P. G. Y. Tsui, "Limitation of CMOS supply-voltage scaling by MOSFET threshold-voltage variation," *IEEE J. Solid-State Circuits,* vol. 30, no. 8, pp. 947–949, Aug. 1995.
[5] T. Kobayashi and T. Sakurai, "Self-adjusting threshold-voltage scheme (SATS) for low-voltage high-speed operation," in *Proc. CICC'94,* May 1994, pp. 271–274.
[6] K. Seta, H. Hara, T. Kuroda, M. Kakumu, and T. Sakurai, "50% active-power saving without speed degradation using standby power reduction (SPR) circuit," in *ISSCC Dig. Tech. Papers,* Feb. 1995, pp. 318–319.
[7] S. Mutoh, T. Douseki, Y. Matsuya, T. Aoki, S. Shigematsu, and J. Yamada, "1-V power supply high-speed digital circuit technology with multithreshold-voltage CMOS," *IEEE J. Solid-State Circuits,* vol. 30, no. 8, pp. 847–854, Aug. 1995.
[8] M. Matsui, H. Hara, K. Seta, Y. Uetani, L.-S. Kim, T. Nagamatsu, T. Shimazawa, S. Mita, G. Otomo, T. Ohto, Y. Watanabe, F. Sano, A. Chiba, K. Matsuda, and T. Sakurai, "200 MHz video compression macrocells using low-swing differential logic," in *ISSCC Dig. Tech. Papers,* Feb. 1994, pp. 76–77.
[9] T. Sakurai and A. R. Newton, "Alpha-power law MOSFET model and its application to CMOS inverter delay and other formulas," *IEEE J. Solid-State Circuits,* vol. 25, no. 2, pp. 584–594, Apr. 1990.

Tadahiro Kuroda (M'88) received the B.S. degree in electronic engineering from the University of Tokyo, Tokyo, Japan, in 1982.

In 1982, he joined Toshiba Corporation, Japan, where he was engaged in the development of CMOS design rules, CMOS gate arrays, and CMOS standard cells. From 1988 to 1990, he was a Visiting Scholar at the University of California, Berkeley, doing research in the field of computer-aided design of VLSI's. In 1990, he was back in Toshiba and involved in the development of BiCMOS ASIC's and ECL gate arrays. In 1993, he joined the Semiconductor Device Engineering Laboratory in Toshiba where he was engaged in the research and development of high-speed circuits for telecommunication. Since 1996, he has been responsible for the research and development of multimedia LSI's including media processors and video compression/decompression LSI's in the System ULSI Engineering Laboratory in Toshiba. His research interests include high-speed, low-power, low-voltage circuit design techniques.

Mr. Kuroda is serving as a program committee member for the VLSI Circuits Symposium. He is a member of the IEEE and the Institute of Electronics, Information, and Communication Engineers of Japan.

Tetsuya Fujita was born in Tokyo, Japan, on August 30, 1963. He received the B.S. degree in electronic engineering from Hosei University, Tokyo, Japan, in 1986.

In 1986, he joined Toshiba Corporation, Kawasaki, Japan, where he was engaged in the establishment of CMOS and ECL gate array libraries. Since 1996, he has been with System ULSI Engineering Laboratory at Toshiba, where he has been involved in the research and development of communication LSI's. His current interests include ATM LSI's and high-speed, low-power, low-voltage techniques in CMOS.
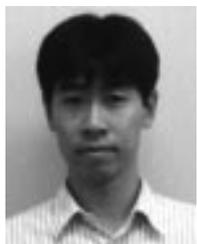
Shinji Mita was born in Aichi, Japan, on March 18, 1970. He received the B.S. degree in electrical engineering from the University of Kyushu, Fukuoka, Japan, in 1992.

In 1992, he joined Toshiba Corporation, Kawasaki, Japan. Since 1992, he has been with Semiconductor Device Engineering Laboratory at Toshiba, where he has been involved in the research and development of multimedia LSI's. His current interests include high-speed, low-power, low-voltage techniques in CMOS.

Tetsu Nagamatsu was born in Yamaguchi, Japan, on August 13, 1960. He received the B.S. degree in applied physics from Waseda University in 1984 and the M.S. degree in energy science from Tokyo Institute of Technology.

He joined the Semiconductor Device Engineering Laboratory, Toshiba Corporation, Kanagawa, Japan, in 1986. He was engaged in the research and development of BiCMOS logic gate, GA and memory macros. Then he was also engaged in the design of DCT/IDCT macro for MPEG2 Decoder. He currently belongs to the System LSI Laboratory. And he is engaged in the development of high-speed CMOS differential circuits for telecommunication area.

Shinichi Yoshioka was born in Tokyo, Japan, in 1963. He received the B.E. and M.E. degrees in electric engineering from Keio University, Tokyo, Japan, in 1987 and 1989, respectively.

In 1989, he joined the Research and Development Center, Toshiba Corporation, Kawasaki, Japan, where he has been working on logic LSI's. Since 1996, he has been engaged in the development of ULSI processors at the System ULSI Laboratory, Toshiba Corporation.

Kojiro Suzuki was born in Kawasaki, Japan, on October 11, 1967. He received the B.S., M.S., and Ph.D. degrees in superconductivity from University of Tokyo, Tokyo, Japan, in 1990, 1992, and 1995, respectively. His Ph.D. work was on design and fabrication of a high-sensitivity SQUID with $Nb/AlO_x/Nb$ Josephson junctions.

In 1995, he joined Toshiba Corporation, Kawasaki, Japan, where he was engaged in the design of low-voltage bus circuits, supply voltage controller, and the research of CMOS low-power design rules. His present interest is low-power, low-voltage techniques in CMOS circuits.

Fumihiko Sano was born in Shiga, Japan, on March 18, 1967. He received the B.S. degree in electrical engineering from Fukui National College of Technology, Japan, in 1988.

He joined Toshiba Microelectronics Corporation, Kawasaki, Japan. He then joined Toshiba's Microelectronics Engineering Laboratory, Kawasaki, Japan, where he has been engaged in the research and development of BiCMOS macrocells for high-performance ASIC's. He has also been engaged in the research and development of VLD macrocells implemented in MPEG2-decoder LSI.

Masayuki Norishima was born in Tokyo, Japan, on January 6, 1962. He received the B.S. degree in pure and applied science from Tokyo University, Tokyo, Japan, in 1986.

He joined the Semiconductor Device Engineering Laboratory, Toshiba Corporation, Kawasaki, Japan, in 1986. From 1986 to 1990, he was engaged in the research and development of process/device technology for Bi-CMOS logic VLSI. From 1990 to 1995 he was engaged in the research and development of process/device technology for high speed CMOS logic VLSI. In September 1995, he joined the LSI Division 2, Toshiba Corporation, Kawasaki, Japan, where he is working on the development of process/device technology for CMOS logic VLSI, focused on mass production in fabs.

Mr. Norishima is a member of the Japan Society of Applied Physics.

Masayuki Murota was born in Kanagawa, Japan. He received B.E. and M.E. degrees in electrical engineering from Hosei University, Tokyo, Japan in 1988 and 1990, respectively.

He joined the Microelectronics Engineering Laboratory, Toshiba Corporation, Kawasaki, Japan, in 1990. He had been engaged in the research and development for process technology of multilevel interconnection of LSI's. Since 1995, he has been engaged in the research and development of high performance microprocessor and logic devices.

Mr. Murota is a member of the Japan Society of Applied Physics.

Makoto Kako was born in Aichi, Japan, in 1972. He graduated from the Technical High School of Higashiyama, Aichi, Japan, in 1991.

He joined the Toshiba Corporation, Kawasaki, Japan, in 1991. From 1991 to 1992, he attended a one-year technical training program at the Toshiba Computer School, Kawasaki, Japan. He has been engaged in the research and development of process technology for high performance microprocessor and logic devices.

Masaaki Kinugawa (M'91) was born in Hyougo, Japan, in 1958. He received the B.S. degree in physics from Kyoto University, Kyoto, Japan, in 1981 and the M.S. degree in applied physics from Tokyo University, Tokyo, Japan in 1983.

In 1983, he joined the Semiconductor Device Engineering Laboratory, Toshiba Corporation, Kawasaki, Japan. He was engaged in the development of CMOS process and device technology for static RAM and high performance RISC chips. He moved to the ULSI Device Engineering Laboratory in 1996.

Mr. Kinugawa is a member of IEEE Electron Device Society.

Masakazu Kakumu (M'90) received the B.S., M.S., and D.E. degrees in electrical engineer from Waseda University, Tokyo, Japan in 1979, 1981, and 1992, respectively.

He joined the Semiconductor Device Engineering Laboratory, Toshiba Corporation, Kawasaki, Japan in 1981. He was engaged in the development of CMOS process and device technology for high-density static RAM. From 1989 to 1990, he worked with Silicon Process Laboratory, Hewlett Packard Company, where he studied on low-temperature CMOS device. He moved to the RISC Processor Engineering Department in 1996. Currently, his responsibilities involve development of CMOS process and device technology for high performance microprocessor and logic devices.

Dr. Kakumu is a member of IEEE Electron Device Society and the Institute of Electronics Information and Communication Engineers.

Takayasu Sakurai (S'77–M'78) received the B.S., M.S., and Ph.D degrees in electronic engineering from University of Tokyo, Tokyo, Japan, in 1976, 1978, and 1981, respectively. His Ph.D. work is on electronic structures of a Si–SiO$_2$ interface.

In 1981, he joined the Semiconductor Device Engineering Laboratory, Toshiba Corporation, Japan, where he was engaged in the research and development of CMOS dynamic RAM and 64 Kb, 256 Kb SRAM, 1 Mb virtual SRAM, cache memories, and BiCMOS ASIC's. During the development, he also worked on the modeling of interconnect capacitance and delay, new memory architectures, hot-carrier resistant circuits, arbiter optimization, gate-level delay modeling, alpha/nth power MOS model and transistor network synthesis. From 1988 through 1990, he was a visiting scholar at University of California, Berkeley, doing research in the field of VLSI CAD. From 1990, back in Toshiba, he managed multimedia LSI development including media processors and video compression/decompression LSI's. From 1996, he is a Professor at the Institute of Industrial Science, University of Tokyo, working on low-power and high-performance LSI designs.

Dr. Sakurai is serving as a program committee member for CICC, DAC, ICCAD, ICVC, ISPLED, ASP-DAC, and FPGA Workshops. He is a technical committee chairperson for the VLSI Circuits Symposium. He is a member of the IEICEJ and the Japan Society of Applied Physics.