MSB-Controlled Inversion Coding for a Low-Power Matrix Transposer

Kyeounsoo Kim and Peter. A. Beerel

This letter proposes a low-overhead MSB-controlled inversion coding technique to reduce the transition activity in a matrix transposer. A family of designs is identified in which this technique is applied to different bit slices of the matrix data and the optimal design within the family is determined using transition activity analysis for DCT and IDCT applications.

Introduction: The increasing demand for portable and wireless multimedia applications that rely on limited battery energy has made low power architectures and designs for these applications critical. Since real-time matrix transposition consumes a large fraction of the power in multi-dimensional image and signal processing, low-power matrix transposers are particularly important.

When a digital circuit is implemented in CMOS, the largest energy is typically consumed by the dynamic power dissipation, which is expressed as $(1/2)TC_LV_{dd}^2f$, where T is the circuit's transition activity, C_L is the total capacitance, V_{dd} and f are the supply voltage and the frequency of operation, respectively. The product TC_L is referred to as the total switching capacitance of the circuit [1]. Various coding schemes, including bus-invert coding, require significant overhead circuitry and are consequently only efficient when the switched capacitance is relatively large, as is the case with system-level buses [2].

This letter explores the on-chip application of a low-overhead adaptation of bus-invert coding that we call most-significant-bit controlled inversion coding (MSB-CIC), to a matrix transposer. The motivation for our approach stems from the observation that typically high fractions of data bits in 2-D DCT/IDCT are sign extension bits (SEB). Thus the key idea is to reduce the overhead by using the most significant data bit as the invert control signal that is applied to a subset of the remaining data bits. Since this approach does not require any extra control signal and circuit for extracting the Hamming distance between the previous and current data as in traditional bus-invert coding strategy [2], it has much lower overhead.

MSB-controlled inversion coding algorithm:
Because MSB-CIC works best for those bits that have high spatial correlation, which for DCT/IDCT data tends to be in the most significant bits, we characterize a family of designs in which the MSB-CIC is applied to the most significant k bits.

Consider a sequence of L-bit data $\boldsymbol{B}_i = (b_i^0, b_i^1, \ldots, b_i^{L-1})$ to be encoded into the sequence $\boldsymbol{C}_i = (c_i^0, c_i^1, \ldots, c_i^{L-1})$, where i is the time index. The MSB-CIC is a mapping function from $\boldsymbol{B}_i$ to $\boldsymbol{C}_i$ for which $c_i^l = b_i^{L-1} \oplus b_i^l$ for $(L-1)\text{-}k \le l < L-1$, and $c_i^l = b_i^l$ otherwise. As mentioned earlier, the different values of k identifies a family of designs for which we expect MSB-CIC to be most effective.

Notice that when the magnitude of $\boldsymbol{B}_i$ is small, the Hamming distance between adjacent $\boldsymbol{C}_i$ data will be smaller than that of adjacent $\boldsymbol{B}_i$ data. Consequently, the idea behind our MSB-CIC algorithm is to reduce the transition activity in the matrix transposer by transposing $\boldsymbol{C}_i$ instead of $\boldsymbol{B}_i$ and then reconstruct $\boldsymbol{B}_i$ after transposition.

Low-power matrix transposer: The proposed low-power matrix transposer is composed of a conventional matrix transposer, additional MSB-CIC and decoding (MSB-CID) circuits as shown in Fig. 1. Most commonly, matrix transposition is implemented using a two-dimensional array of transposition cells (TCs) each containing a multiplexer and a register [3-4]. The conventional matrix transposer is typically constructed with the 2-dimensional array of TCs, and the multiplexers inside the TCs and the output multiplexers route data either vertically or horizontally depending on the status of the row/column control signal, *sw*. For clarity, we decompose the conventional matrix transposer into an upper slice (k^{th} and higher bits) in which the input data is encoded and a lower slice ($k\text{-}1^{th}$ and lower bits) in which the input data is not encoded.

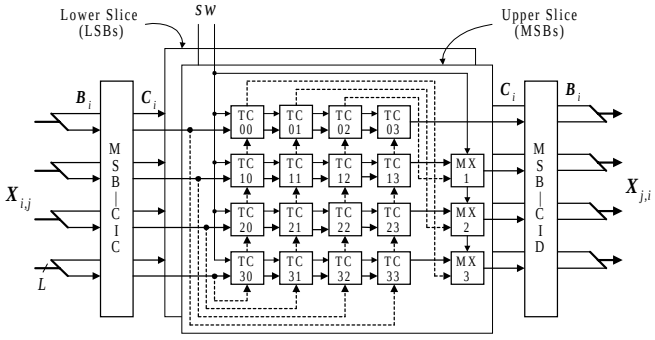

Fig. 1 *Proposed low-power matrix transposer (4×4)*

As shown in Fig. 2, the MSB-CIC and MSB-CID circuits are implemented using a bank of exclusive-OR gates, one gate for each of the k most significant bits, excluding the MSB itself. The MSB is passed directly into the matrix transposer and is the conditional inversion control signal for all other most significant bits. Unlike the conventional conditional inversion coding, MSB-CIC doesn't require an additional control signal for coding and decoding.

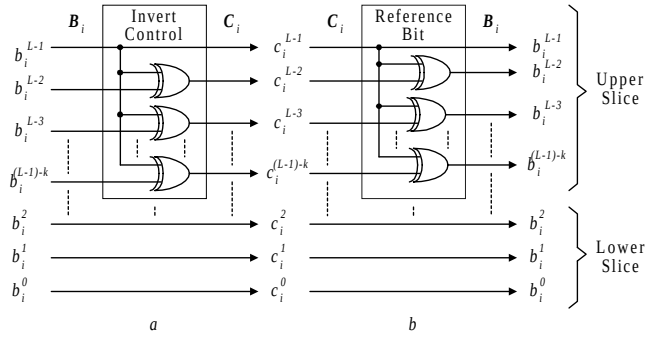

Fig. 2 *Circuits for MSB-controlled inversion coding*
a MSB-CIC circuit
b MSB-CID circuit

Results: We first gathered statistics of the intermediate data to be transposed between two 1-D DCT/IDCT processors for 10 frames of three image sequences (i.e., flower garden, football, and table tennis), each having 720×480 resolution, or equivalently 81,000 8×8 blocks. As illustrated in Fig. 3, we observed that for more than 50% of DCT data, the 8 most significant bits are SEBs and that for more than 80% of IDCT data, all bits are SEBs. The statistics of the data reflects the fact that most of the DCT/IDCT data is small in magnitude for which most bits are SEBs. Therefore, a large fraction of the transition activity can probably be attributed to adjacent data values that alternate between small positive and small negative numbers.

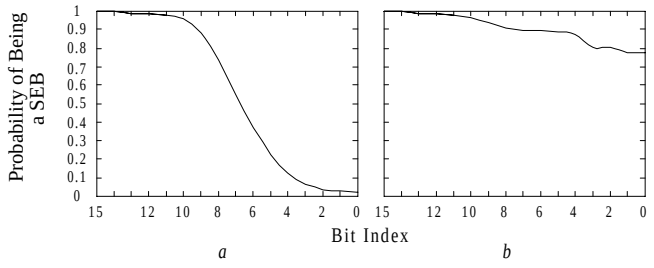

Fig. 3 *Average probability distribution of SEBs*
a DCT data
b IDCT data

We then determined the optimal value of k for the same data through a bit-level hardware model of the conventional and proposed architectures. The transition activity is measured in both the TC array and the MSB-CIC and MSB-CID circuits. It is assumed that the switched capacitance of one TC is twice that of a single exclusive-OR gate. The percentage reduction of the transition activity for the MSB-CIC is measured by $t(\boldsymbol{C})$ - $t(\boldsymbol{O})$, where $t(\boldsymbol{C})$ is the transition reduction ignoring the transition activity of the MSB-CIC and MSB-CID circuits and $t(\boldsymbol{O})$ is the fraction of transition activity of the overhead circuits.

Fig. 4 shows the average reductions with and without including the overhead for all values of k. The results indicates that the optimal value for k for DCT and IDCT data of the matrix transposition is the 9^{th} and 12^{th} bit from MSB, respectively. Although not shown, our experiments suggest that the optimal value of k does not significantly vary across different image sequences. The resulting reduction in transition activity for the optimal value of k is approximately 33% for DCT data and 46% for IDCT data. Due to the randomly distributed bit patterns in the LSBs, as the range of the MSB-CIC exceeds k bits from the MSB, the efficiency is lowered because the reduced transition activity does not overcome the additional overhead.

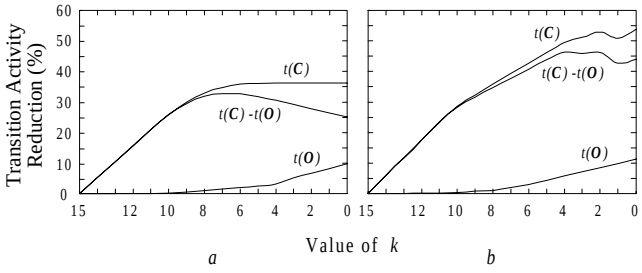

Fig. 4 *Optimal k-bit position extraction from the MSB-CIC*
a DCT data
b IDCT data

Conclusions: We have presented MSB-CIC technique to reduce the transition activity with low hardware overhead, and have shown its on-chip realization for reducing power of matrix transposer. Whereas the bus-invert coding techniques are recommended *only* for system level high capacitive buses [2], the proposed MSB-CIC technique can be successfully applied for on-chip two's complement buses that typically transmit small numbers.

The technique can be viewed as an efficient *partial* and *local* transformation of two's complement data into sign-and-magnitude data. Thus, our matrix transposer design can seamlessly be incorporated into existing typical data-paths that are designed for two's complement data. This feature is important since most existing multimedia standards assume two's complement representation.

Acknowledgement: This research has been partially supported by NSF Grant CCR-9812164, and was done while the first author was on leave from Korea Telecom and he has been partially supported by KOSEF.

References

1 CHANDRAKASAN, A. P., and BRODERSEN, R. W.: 'Minimizing power consumption in digital CMOS circuits', *Proceedings of the IEEE*, 1995, vol. 83, no. 4, pp. 498-523.

2 STAN, M. R., and BURLESON, W. P.: 'Bus-Invert Coding for Low-Power I/O', *IEEE Transactions On Very Large Scale Integration (VLSI) Systems*, 1995, vol. 3, no. 1, pp. 49-58.

3 FATEMI, O., and PANCHANATHAN, S.: 'VLSI Architecture of a Scalable Matrix Transposer', *IEEE International Conference on Innovative systems in Silicon*, 1996, pp. 382-391.

4 TIERNO, J. A., and KUDVA, P.: 'Asynchronous Transpose-Matrix Architectures', *International Conference on Computer Design*, 1997, pp. 423-428.