Improved Memoryless RNS Forward Converter Based on the Periodicity of Residues
A. B. Premkumar, Senior Member, IEEE, E. L. Ang, and Edmund M.-K. Lai, Senior Member, IEEE

Abstract—The residue number system (RNS) is suitable for DSP architectures because of its ability to perform fast carry-free arithmetic. However, this advantage is over-shadowed by the complexity involved in the conversion of numbers between binary and RNS representations. Although the reverse conversion (RNS to binary) is more complex, the forward transformation is not simple either. Most forward converters make use of look-up tables (memory). Recently, a memoryless forward converter architecture for arbitrary moduli sets was proposed by Premkumar in 2002. In this paper, we present an extension to that architecture which results in 44% less hardware for parallel conversion and achieves 43% improvement in speed for serial conversions. It makes use of the periodicity properties of residues obtained using modular exponentiation.

Index Terms—Forward and reverse converters, periodicity property, processing elements, residue number system (RNS).

I. INTRODUCTION

Arithmetic operations based on residue number systems (RNS) can be carried out without intermediate carry digits. However, intermodular operation and conversion between number systems are awkward and have prevented the wide-spread use of RNS. Converting a number from a binary representation to its RNS equivalent is known as forward conversion while the inverse operation is called reverse conversion. Even though reverse conversion is generally more complex, forward conversion for arbitrary moduli sets is not simpler. For special moduli sets of the form, forward converters require only modular adders and therefore can be easily implemented. However, forward conversion for arbitrary moduli sets is memory intensive. There are three main approaches for forward conversion. The first approach involves precomputing all possible values that the conversion requires and storing these values in memory [2], [3]. The second approach involves using efficient arithmetic units called processing elements (PE) along with some memory. In both cases, the memory size requirement increases as the dynamic range increases. The third approach is memoryless in that it involves only combinatorial logic in the design. A framework for memoryless forward conversion has recently been introduced [1].

In this paper, we present an improvement to the memoryless architecture in [1]. The improved design makes use of the cyclic properties of residues to reduce hardware requirement. This paper is organized as follows. Residue numbers and some early forward converters are introduced in Section II. In Section III, the converter proposed in [1] is briefly presented. Periodicity properties of residues are discussed in Section IV. An improved converter which combines the periodicity property with the converter proposed in [1] is then formulated. The advantages of the new converter are discussed in Section VI and comparisons between converters are made.

II. RNS AND ROM-BASED FORWARD CONVERSION

Any n-bit nonnegative integer X in the range $0 \leq X \leq 2^n - 1$ can be represented in the weighted binary system as $X = \sum_{j=0}^{n-1} b_j 2^j$ where $b_j \in \{0, 1\}$. In RNS, X is represented by a set of k residues as $X = \{ r_0, r_1, \ldots, r_{k-1} \}$, where $r_i = X \mod m_i$. In this system, the set $\{ m_i, 0 \leq i < k \}$ constitute the moduli set [4]. The range of the RNS representation is given by $R = \prod_{i=0}^{k-1} m_i - 1$.

Converting a number from its binary representation to an RNS representation is performed by the forward converter. We shall briefly describe some earlier converters as well as a more recent one [1] on which the present work is based.

In Alia and Martinelli’s method [2], the residue corresponding to the jth bit of an n-bit binary number with respect to $m_i$ is determined and stored in a register. Two such registers storing the residues of adjacent bits are combined in a PE. For a given n-bit number, depending on the value of the bit $b_j$, either the register content or a 0 is output. The outputs of two adjacent PEs are combined using a modular $m_i$ adder. This first level of computation requires $(n/2)$ modular adders. The outputs from several adders are combined in a second level of adders and this process is continued until the final residue is obtained. A modification to this method was proposed by them [3] in which the n-bit word is partitioned into $(n/q)$ smaller words. The residues corresponding to each group of q bits are stored in a ROM which are then accessed and added by modular adders. Capocelli and Giancarlo [5] suggested storing the residue corresponding to the first bit in a group of bits, doubling this residue mod $m_i$ and evaluating the residue of the next power of 2 in that group. This process can be pipelined and the residues corresponding to only one group of bits need to be stored. Hence, this results in more efficient use of ROM.

Pieprask, in [6] and subsequently in [7], [8] proposed residue generators based on the periodic properties of residues. His residue converters are based on half and full periods. They make use of two different periods, namely, the period $P(m_i)$ of odd moduli and the half period $HP(m_i)$ of odd moduli. $P(m_i)$ exists for all moduli while $HP(m_i)$ only exists for some odd moduli. Both these periods are calculated using recursion and a table of periods and half periods for various moduli are generated. Using properties associated with periodicity and half periodicity, Pieprask proposed four architectures that are suitable for n input generators modulo $m_i$. Ananda Mohan [9] proposed partitioning the given n-bit number based...
on the cyclic property of $2^j \mod m$: Using the fact that $2^j, 2^{j+6}, 2^{j+26}$ have the same residues due to the periodicity ($l_0$), fields of $l_0$ bits are added. The width of the result is confined to $l_0$ bits by adding the carry bit resulting from the addition to the result. The residue is then determined using techniques given in [2], [3]. This results in similar PEs being used in the transformation algorithm. Ananda Mohan later used these principles to propose residue converters using ROM. Although minor architectural differences exist, the converters discussed so far invariably use ROM. In the next section, a more recent work on forward converters that does not require the use of memory is considered.

III. FORWARD CONVERSION BASED ON MODULAR EXPONENTIATION

In the context of the present work, modular exponentiation of $n$ refers to $(2^n \mod m)$ for some $m$. Modular exponentiation synthesizes the residue of $2^n$ mod $m$ using only combinatorial logic circuits. Hence, it does not need any memory or PE in residue computations. For the sake of completeness, we shall briefly introduce modular exponentiation by means of an example. A complete treatment of modular exponentiation can be found in [1].

Example 1: Determine the residue of $2^{81 \times 92 \times 180} \mod 13$, where $s_i \in \{0,1\}$

$$[2^{81 \times 92 \times 180}]_{13} = [2^{216} + 4^{50} + 2^{50} + 80]_{13} = [216, 4^{50}, 2^{50}]$$

The residue is the cyclic nature of its residues. Some observations are illustrated in Example 2.

1) Case 1: If $s_1 = 0$, $s_2 = 0$, (1) reduces to 1, 3, 9, and 1 for $s_3 = s_2 = 0$; $s_3 = 0$, $s_2 = 1$; $s_3 = 1$, $s_2 = 0$; and for $s_3 = s_2 = 1$: respectively. A function $g_0$ can be synthesized for different combinations of $s_3$ and $s_2$ as shown in

$$g_0 = s_3 \overline{s}_2 + 0 + 2s_3s_2 + 1.$$ (2)

When the function $g_0$ is evaluated for different combinations of $s_3$ and $s_2$ with $s_1 = s_0 = 0$, we have $\{1, 3, 9, 1\}$. $g_0$ is the cyclic nature of its residues. Some observations are illustrated in Example 2.

2) Case 2: If $s_1 = 0$ and $s_2 = 1$, (1) yields $\{2, 6, 18\}$ for different combinations of $s_3$ and $s_2$. A second function $g_1$ can be obtained as follows:

$$g_1 = 0 + 4(s_2 \oplus s_3) + 2(s_3 \oplus s_2) + 3s_3s_2.$$ (3)

3) Case 3: Similarly, function $g_2$ can be obtained for $s_1 = 1$ and $s_0 = 0$

$$g_2 = 8(s_2 \oplus s_3) + 4(s_2 \oplus s_3) + 2s_3s_2 + 0.$$ (4)

4) Case 4: Function $g_3$ is obtained when $s_2 = s_0 = 1$

$$g_3 = 8(s_3 \oplus s_2) + 4s_3s_2 + 2(s_3 \oplus s_2) + 3s_3s_2.$$ (5)

In all of the above functions, $\oplus$ denotes the bitwise exclusive-or logical operation. These functions use the positional weights of the binary number system in arriving at the values for different $s_3$ and $s_2$. Hence, they can be written as follows:

$$g_0 = \{s_3s_2, 0, \overline{s}_3s_2, 1\}$$

$$g_1 = \{0, (s_2 \oplus s_3), (s_3 \oplus s_2), s_3s_2\}$$

$$g_2 = \{s_2 \oplus s_3, (s_2 \oplus s_3), (s_3 \oplus s_2), 0\}$$

$$g_3 = \{s_2, 0 \oplus s_3, (s_3 \oplus s_2), s_3s_2\}.$$ (9)

Forward conversion from binary to residues can be accomplished by determining the residue of each nonzero $2^{81 \times 92 \times 180}$ term of the given binary number.

The conversion can either be performed on each nonzero bit sequentially or on all bits simultaneously in parallel. Trade-offs can be made between hardware cost and delay by using a combination of serial and parallel methods. Hardware complexity for the parallel implementation mainly arises from the multiplexers since the gates used to generate the MIN terms in obtaining $g_i$ can be shared among different exponents within the same modulus as well as among different moduli.

IV. PERIODICITY PROPERTIES OF RESIDUES

An interesting property of determining the residue of $2^n \mod m$ is the cyclic nature of its residues. Some observations are made about the periodicity properties of these residues.

1) For certain odd moduli $m_k$, there exists a period given by $m_k = 1$ after which the residues repeat. Consider $m_1 = 5$. The residues of $2^n \mod 5$ for different $n$ are $1, 2, 3, 4, 1, 2, 3, 4, 1, 2, 3, 4, 1$. The periodicity of the residues is $5 = 1 = 4$. Whenever the period is $m_k = 1$, it is referred to as basic period.

2) There exists for certain other odd moduli $m_j$, a period that is shorter than the basic period after which the residues repeat.

3) In the case of even moduli other than those that belong to the $2^n$ family, a short period exists after an initial subset of residues [10], [11]. These three properties are illustrated in Example 2.

Example 2: Consider $m_2 = 9$. The residues of $2^n \mod 5$ for different $n$ are given by the set $\{1, 2, 4, 8, 7, 5, 1, 2, 4, 8, 7, 5, 1\}$. The residues repeat after $n = 5$. In this case, there exists a period that is shorter than the basic period, namely, 6. Now, consider $m_3 = 10$. The corresponding residue set in this case is $\{1, 2, 4, 8, 6, 2, 4, 8, 6, 1, 2, 4, 8, 6, 2, 4, 8, 6\}$. It is observed that there is a short period of 4 after the first residue, 1.

V. FORWARD CONVERSION BASED ON PERIODICITY PROPERTIES

In this section, we propose an improved forward converter using modular exponentiation and periodicity property of the residues discussed in the previous section. The proposed converter differs from other converters in that it does not use any PEs but is based on the converter proposed in [1].

In the case of odd moduli, the residues have either a basic period or a short period. The given binary number is partitioned based on either of these periods. Without loss of generality, either the basic or the short period can be assumed to be $p$. Hence, $2^{p}, 2^{p}, 2^{2p}, \ldots$ will all have the same residues. The partitioned fields are added using $p$-bit adders. The length of the result is confined to $p$ bits by adding the carry bit back to the result. This is illustrated by the following example.

Example 3: Consider the following 32-bit number:

$$0001100100011101110001111100.$$
The decimal equivalent of this number is 421,185,084 and its residue with respect to modulus 23 is 22. From [9, Table II], the periodicity of the residues is 11. Therefore, the number is partitioned into three 11-bit fields and added as shown below. A zero has been appended at the MSB position to make the number 33 bits to facilitate easy partitioning into 11-bit fields.

\[
\begin{array}{cccccccccccccc}
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\
\end{array}
\]

1 0 0 1 1 1 1 1 1 1 0 0.

In the above, the top row corresponds to the most significant 11-bit field of the given binary number while the third row corresponds to the least significant 11-bit field. The last row is the result obtained by adding the three rows. In the addition, the carry of 1 is added to the LSB of the result so that the length of the final result is 11 bits. Forward conversion is now performed on the result using modular exponentiation on only 11 bits as opposed to all 32 bits as proposed in [1]. The residue computation of \( m \mod 2^p \) can be applied sequentially or in parallel. This is discussed in the next section along with a comparison between [1, 9] and the proposed converter.

It should be noted that even moduli frequently occur in the use of residue number applications. One such moduli set is \( \{2m - 1, 2m, 2m + 1\} \). Hence, there is a need for forward conversion for even moduli. Residue computation in this case can be performed using the following theorem.

**Theorem 1:** Let \( m \) be an even integer that is not a power of 2. It can be expressed as \( m = 2^p \times y \) where \( p \) is a positive integer and \( y \) is an odd number. The residue of an integer \( X \) modulo \( m \) can be expressed as

\[
[X]_m = \left\lfloor \frac{X}{2^p} \right\rfloor_y \times 2^p + [X]_{2^p}.
\]

**Proof:** \( X = QP + R \) where \( Q \) is the quotient obtained by dividing \( X \) by \( P \) and \( R \) is the residue. Let \( P \) be even and composite not belonging to \( 2^m \) set. Let \( P = y \times 2^p \). Then

\[
R = X \mod P.
\]

\( X \) can also be written as

\[
X = Q_1 \times 2^p + R_1
\]

where \( Q_1 \) is the quotient and \( R_1 \) is the residue obtained dividing \( X \) by \( 2^p \)

\[
0 \leq R_1 < 2^p.
\]

In this case

\[
R_1 = [X]_{2^p}.
\]

However, \( Q_1 \) can be written as

\[
Q_1 = Q_2 \times y + R_2.
\]

\( Q_2 \) is the quotient and \( R_2 \) is the remainder obtained by dividing \( Q_1 \) by \( y \)

\[
0 \leq R_2 < y.
\]

But \( Q_2 = \lfloor (X)/(2^p) \rfloor_y \) from (12). Since \( R_2 \) is the residue of \( Q_1 \)

\[
R_2 = [Q_1]_y = \left\lfloor \frac{X}{2^p} \right\rfloor_y.
\]

We know that

\[
X = QP + R = Q_1 \times 2^p + R_1.
\]

From (13) and (14)

\[
QP + R = (Q_2 \times y + R_2) \times 2^p + [X]_{2^p}
\]

\[
= Q_2 \times y \times 2^p + R_2 \times 2^p + [X]_{2^p}
\]

\[
= Q_2 \times P + \left\lfloor \frac{X}{2^p} \right\rfloor_y \times 2^p + [X]_{2^p}.
\]

Comparing \( X = QP + R \) with (15), we have

\[
R = \left\lfloor \frac{X}{2^p} \right\rfloor_y \times 2^p + [X]_{2^p}.
\]

Note that \([X]_{2^p}\) can be obtained by simply considering the least significant \( p \) bits of \( X \). The first term in (10) is the \( n - p \) most significant bits of the \( X \), where \( X \) is an \( n \)-bit number. Evaluation of \( \lfloor (X)/(2^p) \rfloor_y \) follows the procedure for odd modulus once \( y \) is odd. This is illustrated by the following example.

**Example 4:** To determine the residue of the same 32-bit number with respect to 24, the modulus is written as a composite number as \( 24 = 2^3 \times 3 \). The residue \([X]_{24}\) is simply the three LSB and is therefore, 4. The remaining 29 bits are partitioned into 2-bit fields since the periodicity of 3 is 2 [9, Table II]. The two-bit fields are then added and the result is confined to 2 bits by adding the carry back to the result. The modular exponentiation is then performed and the residue is found to be 1. Using (10), the final residue works out to be 12.

**VI. COMPARATIVE ANALYSIS**

Most digital signal processing (DSP) algorithms can be conveniently executed using a wordlength of 32 bits. A suitable moduli set would be \( \{3, 5, 7, 11, 13, 17, 19, 23, 31\} \). Architectures using modular arithmetic need to implement circuits that generate functions such as those given by (6) to (9). Note that there are several terms common to \( g_k \) functions in these equations. Consequently, the circuit implementing these functions can be derived from commonly occurring terms. The common terms can even be shared among different moduli in the given set. In the implementation, the number of multiplexers required depends on whether serial or parallel technique is employed. The number of gates required for both serial and parallel techniques for our proposed converter based on this moduli set is estimated and compared.
with that proposed in [1] and [9]. Figs. 1 and 2 show examples of the serial and parallel implementations of the proposed method. These figures are specific for Example 1. Reference [7, Table IV] discusses the number of gates that are required and delay issues with regard to specific moduli sets such as $2^n - 1$ and $2^n - 1 + 1$ that do not use memory. The other entries in [7, Table IV] discuss comparison for arbitrary moduli sets. However, these converters use ROM and hence are not considered for comparison here. Since the proposed forward converter is for arbitrary moduli sets that do not use memory, the proposed converter complexity has been compared with that in [1], [9]. Further, hardware and speed are two parameters that are used in the comparison and both these parameters are for modulus 19 only since this modulus has the longest period, namely, 18.

A. Serial Method

For comparison with [1], a worst case scenario when all bits in the given number are nonzero is considered. In Tables I and II, the second column explains how the conversion is accomplished. The exponent 5 in column 2 indicates that there are 32 bits in the binary number. Out of 5 bits, three MSB bits are used to generate $g_i$ functions. The two LSB bits are used in multiplexing these functions.

These are clearly shown in the case of a 4-bit exponent in Example 1. Since each term in $g_i$ is accessed serially and residue computation for each bit is performed taking one bit at a time, a single multiplexer is sufficient. Hence, for a 32-bit number, the residue computation is repeated a maximum of 32 times. Each computation requires 5 clock cycles since there are 5 bits in each residue and each bit in the residue is output sequentially. A total of 132 nand gates is required for generating the $g_i$ functions. All 32 5-bit residues are added using a single modular adder to determine final residue.

The proposed converter requires the given decimal number to be partitioned into fields based on the periodicity and then added. In the chosen set of moduli, a maximum period of 18 occurs for modulus 19 [9]. Residue computation is performed 18 times and the 5-bit residues are added using a single modulo 19 adder. Each of the modulo 19 adder uses 5 Full Adders (FA), 5 D Flip Flops (FF) and two numbers of five 2-to-1 multiplexers and 5 inverters. The serial method uses one 4-to-1 multiplexer which is built using three 2-to-1 multiplexers. Hence, the total number of 2-to-1 multiplexers required is 13. Each 2-to-1 multiplexer requires 9 nand gates and hence a total of 117 nand gates is required for multiplexers. [1] also requires a modulo 19 adder and a 4-to-1 multiplexer. Hence, the number of nand gates required in this method is the same as that used in the proposed method. [9] uses memory access for computing residues and for the chosen modulus 19, one needs 19 look ups or ROM locations only. This is based on the assumption that a residue is read from the ROM each time an access is made. The addition time using a carry propagation adder is the 10 full adder delay and this is common to all three methods. Table I compares the different architectures.

The proposed method is faster by about 43%. However, if we consider a more moderate distribution of ones and zeros in the given number, the increase in speed may not be as high as 43%, but still better than what is achieved in [1]. Since the number of gates in both methods are more or less the same, no comparison between the two methods in terms of hardware or area is made.

B. Parallel Method

In this method residue computation is performed on all bits simultaneously and all terms required in $g_i$ are accessed in par-
TABLE I

<table>
<thead>
<tr>
<th>Method</th>
<th>No. bits in exponent</th>
<th>No. of 2 to 1 MUX</th>
<th>No. of Gates used in MUX</th>
<th>Total No. of Gates (MUX+Logic+Inverter + D Flip Flops)</th>
<th>Total No. of clock cycles (Max)</th>
<th>No. of FAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ref [1]</td>
<td>5 bits (4 to 1 MUX)</td>
<td>13</td>
<td>117 (includes MUX in adders)</td>
<td>117+132+5+3FF</td>
<td>32 × 5</td>
<td>5</td>
</tr>
<tr>
<td>Proposed Method</td>
<td>5 bits (4 to 1 MUX)</td>
<td>13</td>
<td>117 (includes MUX in adders)</td>
<td>117+132+5+3FF</td>
<td>18 × 5</td>
<td>5</td>
</tr>
<tr>
<td>Ref [9]</td>
<td>5 bits (4 to 1 MUX)</td>
<td>10</td>
<td>90 (includes MUX in adders)</td>
<td>ROM+5INV+5FF</td>
<td>18</td>
<td>5</td>
</tr>
</tbody>
</table>

TABLE II

<table>
<thead>
<tr>
<th>Method</th>
<th>No. bits in exponent</th>
<th>No. of 2 to 1 MUX</th>
<th>No. of Gates used in MUX</th>
<th>Total No. of Gates (MUX+Logic+Inverter + D Flip Flops)</th>
<th>No. of FAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ref [1]</td>
<td>5 bits (4 to 1 MUX)</td>
<td>790</td>
<td>7110 (includes MUX in adders)</td>
<td>7110+132+155+155FF</td>
<td>155</td>
</tr>
<tr>
<td>Proposed Method</td>
<td>5 bits (4 to 1 MUX)</td>
<td>450</td>
<td>4050 (includes MUX in adders)</td>
<td>4050+132+85+85FF</td>
<td>85</td>
</tr>
<tr>
<td>Ref [9]</td>
<td>170</td>
<td>1530</td>
<td>1530 (includes MUX in adders)</td>
<td>ROM+1530+85INV+85FF</td>
<td>85</td>
</tr>
</tbody>
</table>

VII. CONCLUSION

Forward conversions in RNS have been traditionally implemented using lookup tables. Modifications to memory based systems mainly involved using PEs. Although the use of PEs reduces memory requirements, they have not completely eliminated the use of memory. Furthermore, memory based converters require reprogramming for different moduli sets. In this paper we have presented an extension to memoryless binary to RNS converters that makes them less hardware intensive. The method makes use of the periodicity properties of residues obtained using modular exponentiation. As a result, the new converter requires 40% less hardware for parallel conversion and achieves 43% improvement in speed for serial conversions.

ACKNOWLEDGMENT

The authors are indebted to Dr. P. V. Ananda Mohan for his valuable comments and clarifications during the course of this work.

REFERENCES

Improved Memoryless RNS Forward Converter Based on the Periodicity of Residues

Premkumar, A. B.

2006-02-01

http://hdl.handle.net/10179/9618
28/11/2018 - Downloaded from MASSEY RESEARCH ONLINE