Research Article

A Test Data Compression Scheme Based on Irrational Numbers Stored Coding

Hai-feng Wu, 1 Yu-sheng Cheng, 1 Wen-fa Zhan, 2 Yi-fei Cheng, 1 Qiong Wu, 2 and Shi-juan Zhu 1

1 School of Computer and Information, Anqing Normal University, Anqing 246011, China
2 Department of Science Research, Anqing Normal University, Anqing 246011, China

Correspondence should be addressed to Hai-feng Wu; hfeng_wu@163.com

Received 27 June 2014; Revised 2 August 2014; Accepted 2 August 2014; Published 28 August 2014

Academic Editor: Yunqiang Yin

Copyright © 2014 Hai-feng Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Test question has already become an important factor to restrict the development of integrated circuit industry. A new test data compression scheme, namely irrational numbers stored (INS), is presented. To achieve the goal of compress test data efficiently, test data is converted into floating-point numbers, stored in the form of irrational numbers. The algorithm of converting floating-point number to irrational number precisely is given. Experimental results for some ISCAS 89 benchmarks show that the compression effect of proposed scheme is better than the coding methods such as FDR, AARLC, INDC, FAVLC, and VRL.

1. Introduction

According to Moore's law, the integration level of the microchips doubles every 18 to 24 months. As a result, the volume of test data increased dramatically; the test costs become higher; the contradictions between test efficiency and test quality are sharpening. The test data compression technology comes into being to solve the above problems.

Coding compression technology has been widely used because it has such advantages as simple encoding and decompression structure which is independent of the tested chip. A certain encoding scheme is used to encode the original test set. The length of the encoded data is less than the original test set, so as to reduce the test data.

The test set is divided into the sequences of specific law, which can be replaced with new codeword generated by some kinds of coding method. According to the change rule of the length from the original sequence to the codeword, coding methods can be divided into four categories. The first category is fixed-to-fixed coding method, such as dictionary code [1]. The second category is fixed-to-variable coding method, such as Huffman code [2], 9C code [3]. The third category is variable-to-fixed encoding method, such as run-length code [4]. The fourth category is variable-to-variable coding method, such as Golomb code [5], FDR code [6], EFDR code [7], and alternation and run-length code [8].

Among them, the compression method of the first category is the most simple one, but it has the lowest compression efficiency. The compression efficiency of the fourth category is high, but its hardware overhead is larger. The second and third categories have good applicability, which are between the first and fourth categories in terms of compression method, compression efficiency, hardware overhead, and so forth.

It is a new original method to use irrational number to compress test data. It is creative. A scheme of test data compression based on irrational number dictionary code [9] is presented. The encoding rule of this scheme is simple, and the do not-care bit need not be assigned, but it takes extra storage space to store the data dictionary additionally.

A new test data compression scheme based on irrational numbers stored (INS) is presented. It is a fixed-to-variable coding method, which has simple encoding rule and can obtain good compression effect. In this scheme, the test set is converted into floating-point numbers firstly, and then the floating-point numbers are converted into irrational numbers in form of $\sqrt{x}$ (where the numbers $x$ and $r$ are integers) by the successive approximation method. So the storage of the test set can be translated into the storage of
The Scientific World Journal

radicand (integer) and root number (integer). The minimum of corresponding radicand and root number can be found with faster convergence speed. Better compression effect can be obtained when using the INS coding scheme.

The organization of the paper is as follows. Section 2 explains the algorithm of the proposed scheme and gives an example. Section 3 proves the feasibility of the algorithm of the proposed scheme theoretically. The structure of decompressor is presented in Section 4. Section 5 reports the experimental results and analyzes the compression ratio theoretically. Finally Section 6 concludes the paper.

2. INS Coding

In this section, the algorithm of INS coding is described firstly. Then, the flowchart of the scheme is given. Lastly, an example is provided.

2.1. Encoding Rule. The concrete steps of INS coding are as follows.

(1) At first, generate the determinate complete test set named T by automatic test pattern generation tools.
(2) Then cascade the entire test vector, and connect the head of a vector to the tail of previous vector, remembered as S.
(3) Take the first N bits of the test set, and then convert them into a hexadecimal number every four bits a group according to the rules shown in Table 1. Add a decimal point after the first digit, and a hexadecimal floating-point number named f is formed.
(4) Calculate the value of x and r in the formula $\sqrt{x} = f$, by dichotomy to successive approximate f. In this scheme, if the first several bits of f and the root of $\sqrt{x}$ are exactly the same, that is to say, the two are approximately equal to, then consider the following. (I) Calculate $f^2$ first; set bot = $[f^2]$, top = bot + 1, $r = 2$. (II) Calculate $\sqrt{\text{bot}}$, if its value is equal to f, note x = bot, r, and then go to step (5). Otherwise, calculate $\sqrt{\text{top}}$, if its value is equal to f, note x = top, r, and then go to step (5). (III) Consider bot = [bot * $\sqrt{\text{bot}}$], top = [top * $\sqrt{\text{top}}$], and $r = r + 1$. (IV) Set mid = [(bot + top)/2], calculate $\sqrt{\text{mid}}$, if its value is equal to f, note x = mid, r, and then go to step (5). If its value is greater than f, top = mid − 1. If its value is less than f, bot = mid + 1. Repeat step (IV) until bot > top, and then go to step (V). If top is less than mid, set bot = top, top = mid. Otherwise, set top = bot, bot = mid. Repeat step (4) until $\sqrt{x} = f$ is valid and go to step (5). In this step, the median (mid) is always integer in operation process. That can reduce the computation complexity, save the running time, and accelerate the operation process.
(5) Encode x and r in the form of CEBM [10]. Remove the first N bits of S and repeat steps (3) and (4), until S is empty.

Figure 1 shows the detailed flowchart of the scheme.

CEBM is shown in Table 2. Because $x \geq 2$, $r \geq 2$, the run-length starts from 2. The first column is the length of runs and the second column is the number of group. The third column is odd bits of the codeword and the fourth column is even bits. The last column is the corresponding codeword. As can be seen, CEBM has the following characteristics. (1) The length of the odd bits and the even bits is equal to every codeword. (2) The odd bits show the run-length. Add data “1” before the odd bits; the new odd bits are just the corresponding binary number of the run-length. The even bits show the end of the codeword. The codeword continues if the even bit is 0 and finishes if it is 1. (3) The length of codeword increases by two bits from group $A_i$ to group $A_{i+1}$; both the odd bits and the even bits increase one bit. (4) Group $A_i$ contains $2^i$ elements.

(5) The corresponding relationship between the run-length L and group $A_i$ is like $l = \lceil \log_2(L + 1) \rceil - 1$.

For an example, the codeword of 9 is 000011. The odd bits are 001 and the even bits are 001. The even bit needs to be monitored when decoding. If the even bit is 0, it means that codeword continues. If the even bit is 1, it means that codeword finishes. Add data “1” before the odd bits (001); we can get the data 1001, whose corresponding decimal value is 9. CEBM is widely used because of decoding simply and small hardware overhead.

Through the analysis, the scheme has the following three characteristics. (1) The root number is calculated from 2 and increases successively. It can guarantee that the radicand and root number found are minimum. (2) The lower bound and upper bound of the radicand's interval positioned are relatively accurate. At the same time, the binary search method can reduce the time complexity. (3) In the binary search, the medians of the operation process are all integers. It can reduce the computation complexity, save the running time, and accelerate the process of operation.

2.2. Encoding Example. An example is provided to make the scheme clear. Without loss of generality, set the original test set $T = \{00011010, 11101000, 10011111, 11010101, 01011010, 10011001, 01010011, 11101001, \ldots\}$. Cascade all the test vector, and then divide it into sequences of 48 bits. The data flow is 000110101110100010111101101011001010110101100110110100111 11101001, ..., its first 48 bits can convert into a hexadecimal floating-point $f = 1.AE89F995AD3$. (I) Calculate $f^2 = 2.D413CCCF751FCA6F909E9$, bot = $[f^2] = 2$, top = bot + 1 = 3, $r = 2$. (II) Calculate $\sqrt{\text{top}} = \sqrt{3} = 1.BB67A8E8584C$, and go to step (III) because its value is not equal to $f$. (III) Consider bot = [bot * $\sqrt{\text{bot}}$], top = [top * $\sqrt{\text{top}}$], and $r = r + 1$. Repeat step (IV) until bot > top, and then go to step (V). If top is less than mid, set bot = top, top = mid. Otherwise, set top = bot, bot = mid. Repeat step (4) until $\sqrt{x} = f$ is valid and go to step (5). In this step, the median (mid) is always integer in operation process. That can reduce the computation complexity, save the running time, and accelerate the process of encoding. If its value is greater than f, set top = mid − 1. If its value is less than f, set bot = mid + 1. Repeat step (IV) until bot > top (bot = 5, top = 4), and go to step (V). If top is less than mid, set bot = top, top = mid. Otherwise, set top = bot, bot = mid. Then bot = 4, top = 5. Repeat steps (III), (IV), and (V) and we can get the results.
Begin

Generate the test set named $T$, cascade all the vector, remember as $S$

If $S$ is not empty

Take the first $N$ bits, convert into a floating point named $f$

Set bot = $\lfloor f^2 \rfloor$, top = bot + 1, $r = 2$, flag = 0

If $\sqrt{\text{bot}} = f$

Yes

Flag = 1, $x = \text{bot}$

No

If $\sqrt{\text{top}} = f$

Yes

Flag = 1, $x = \text{top}$

No

Flag = 0

If Flag = 0 and Bot ≤ top

Yes

Set mid = $\lfloor (\text{bot} + \text{top})/2 \rfloor$

No

If $\sqrt{\text{mid}} = f$

Yes

Flag = 1, $x = \text{mid}$

No

If $\sqrt{\text{mid}} > f$

Yes

Top = mid - 1

No

Bot = mid + 1

If top < mid

No

Bot = top, top = mid

Yes

Top = bot, bot = mid

Bot = $[\text{bot}^{\sqrt{\text{bot}}}]$, top = $[\text{top}^{\sqrt{\text{top}}}]$, $r = r + 1$

Encode $x$ and $r$, remove the first $N$ bits of $S$

End

Figure 1: Flowchart of INS coding.
Table 1: Test data transformation rules.

<table>
<thead>
<tr>
<th>Test data</th>
<th>Hexadecimal number</th>
<th>Test data</th>
<th>Hexadecimal number</th>
<th>Test data</th>
<th>Hexadecimal number</th>
<th>Test data</th>
<th>Hexadecimal number</th>
</tr>
</thead>
<tbody>
<tr>
<td>0000</td>
<td>0</td>
<td>0100</td>
<td>4</td>
<td>1000</td>
<td>8</td>
<td>1100</td>
<td>C</td>
</tr>
<tr>
<td>0001</td>
<td>1</td>
<td>0101</td>
<td>5</td>
<td>1001</td>
<td>9</td>
<td>1101</td>
<td>D</td>
</tr>
<tr>
<td>0010</td>
<td>2</td>
<td>0110</td>
<td>6</td>
<td>1010</td>
<td>A</td>
<td>1110</td>
<td>E</td>
</tr>
<tr>
<td>0011</td>
<td>3</td>
<td>0111</td>
<td>7</td>
<td>1111</td>
<td>B</td>
<td>1111</td>
<td>F</td>
</tr>
</tbody>
</table>

Table 2: CEBM.

<table>
<thead>
<tr>
<th>Run-length</th>
<th>Group</th>
<th>Odd bits</th>
<th>Even bits</th>
<th>Codeword</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>A₁</td>
<td>0</td>
<td>1</td>
<td>01</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>1</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>00</td>
<td>01</td>
<td>00 01</td>
</tr>
<tr>
<td>5</td>
<td>A₂</td>
<td>01</td>
<td>01</td>
<td>00 11</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>10</td>
<td>01</td>
<td>10 01</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>11</td>
<td>01</td>
<td>10 11</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>00</td>
<td>01</td>
<td>00 00 01</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>00</td>
<td>01</td>
<td>00 00 11</td>
</tr>
<tr>
<td>...</td>
<td>A₃</td>
<td>:</td>
<td>:</td>
<td>:</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>110</td>
<td>00</td>
<td>10 10 01</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>111</td>
<td>00</td>
<td>10 11 11</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td>:</td>
<td>:</td>
<td>:</td>
</tr>
</tbody>
</table>

Therefore, we only need to prove that \( \sqrt{\text{bot}} \) is infinitely close to \( \sqrt{\text{top}} \) when \( r \to \infty \). That is to prove

\[
\lim_{r \to \infty} \frac{\sqrt{|f^r| + 1}}{\sqrt{|f^r|}} = 1
\]

\[
\lim_{r \to \infty} \frac{\ln(\sqrt{|f^r| + 1})}{\ln|f^r|} = \lim_{r \to \infty} \frac{\ln(1 + 1/|f^r|)}{r} = 0
\]

That is to say, we can always find positive integers \( x \) and \( r \) to make \( \sqrt{x} \) and \( f \) to be approximately equal when \( r \) is increasing.

4. Structure of Decompressor

In this section, the structure and basic operation principle of the decompressor for the INS coding scheme are illustrated.

Figure 2 shows the structure of decompressor, which contains a finite state machine, a T flip-flop, a special \( k + 1 \) bits counter and some combinational logic, and the CPU module on SoC chip.
This special $k + 1$ bits counter is unique. Set its lowest bit as 1; move this bit to high with other data when the data needed to be decoded is moved to the counter. The function can be realized by a combination of some simple circuit, and the hardware overhead will not increase significantly. In the $k + 1$ bits counter, the value of $k$ is decided by the maximum of the radicand and the boot number (named as $L_{\text{max}}$) in compression result, $k = \lceil \log_2(L_{\text{max}} + 1) \rceil - 1$.

The root operation can be realized by the CPU module on SoC chip as long as the floating-point processing unit $\times 87$ FPU integrated in CPU. $\times 87$ FPU has its own instruction system, including the commonly used instruction types: floating-point move instructions, floating-point arithmetic operation instructions, floating-point transcendental function instructions, floating-point comparison instructions, and FPU control instructions [11]. Floating-point transcendental function instructions can realize exponent operation (F2XM1 instruction) and logarithmic operation (FYL2X instruction). $\sqrt{x}$ can be converted into $2^{\log_2(x)/2}$, which can be calculated by exponent operation instruction and logarithmic operation instruction.

Decompressor works as follows. (1) First, the FSM makes an enable signal named $en$ into high level, and the encoded data named $bit_{\text{in}}$ is split into two parts: odd bits and even bits by the position of data. That is implemented by the T flip-flop. Odd bits are directly shifted into the $k + 1$ bits counter, and even bits are shifted into FSM. (2) Then, FSM repeats to read the encoded data until the value of even bit equals 1. Meanwhile, the odd bit is shifted into the $k + 1$ bits counter. (3) The data (named $x$) of the $k + 1$ bits counter is shifted into CPU after all odd bits of data are received in the special $k + 1$ bits counter. Repeat step (2); the data of the $k + 1$ bits counter (named $r$) is shifted into CPU. (4) $\sqrt{x}$ is calculated by floating-point processing unit $\times 87$ FPU in CPU, and the first $N$ bits of its binary form are outputted.

5. Experimental Results

In this section, the effectiveness of the INS coding scheme is verified by using experimental results.

INS coding is applied to the MinTest test sets of the largest ISCAS 89 benchmark circuits. The experimental results are shown in Table 3. The first column shows the circuit name, the second column shows the total number of the data in the original test sets, the third column shows the total number of the compressed data, and the fourth column shows the compression ratio.

In order to verify the validity of this scheme, compare with the similar algorithm, as shown in Table 4. Among them, the first column shows the circuit name, the second to sixth columns, respectively, show the compression effect of FDR code [6], INDC code [9], AARLC code [8], FAVLC code [12], and VRL code [13], and the seventh column shows compression effect of the scheme presented. Data from Table 3 shows this scheme besides lower compression ratio in the fifth circuit, and the rest of the circuit has higher compression ratio. And this scheme has a good compression effect on the whole. The average compression ratio reaches 65.79%, which is 20.46%, 9.42%, 7.36%, 5.36%, and 1.85% higher than those of FDR, RLR, AARLC, FAVLC, and VRL.

In order to further demonstrate the effectiveness, the results of proposed scheme are compared with other schemes under the same circumstances by the hard fault set of some ISCAS 89 benchmark circuits. The results as shown in Table 5. As can be seen from Table 4, the overall effect of proposed scheme is better. The average compression ratio of the proposed scheme is 3.87%, 3.82%, and 3.31% higher than those of FDR, AR [14], and ARL [15]. These data show the effectiveness of INS Code.

The experimental results indicate the following. (1) Regardless of the fact that the determined bit of the test data is 0 or 1, it has a little effect on compression ratio in the process of finding the irrational number. (2) Do not-care bit can accelerate the search speed of irrational number, reduce the value of irrational number used to store the test data, and improve the compression ratio.

The influence of the do not-care bit probability on compression is explored. The length of the test data fragment is set to be $K$, and the probability of do not-care bit is set to be $p$. Their relationship can be expressed in the following formula:

$$\sqrt{x} - f_i < 10^{-(K/4)^{i+1}}, \quad i = 1, 2, \ldots, 2^K p,$$

where, $x$, $r$, and $K$ are integers and $f_i$ is a floating-point number.

The compression gain can be calculated by the following formula:

$$\beta = \frac{K}{1 + \log_2(x + 1) + \log_2(r + 1) + \log_2(K + 1)}.$$

<table>
<thead>
<tr>
<th>Circuit</th>
<th>Size of TD bit</th>
<th>Scheme presented</th>
<th>Size/bit</th>
<th>α%</th>
</tr>
</thead>
<tbody>
<tr>
<td>S3578</td>
<td>23754</td>
<td>9868</td>
<td>58.46</td>
<td></td>
</tr>
<tr>
<td>S9234</td>
<td>39273</td>
<td>16157</td>
<td>58.86</td>
<td></td>
</tr>
<tr>
<td>S13207</td>
<td>165200</td>
<td>24998</td>
<td>84.87</td>
<td></td>
</tr>
<tr>
<td>S15850</td>
<td>76986</td>
<td>22555</td>
<td>70.70</td>
<td></td>
</tr>
<tr>
<td>S38417</td>
<td>164736</td>
<td>81611</td>
<td>50.46</td>
<td></td>
</tr>
<tr>
<td>S38584</td>
<td>199104</td>
<td>56981</td>
<td>71.38</td>
<td></td>
</tr>
<tr>
<td>Avg.</td>
<td></td>
<td></td>
<td></td>
<td>65.79</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Circuit</th>
<th>FDR</th>
<th>AARLC</th>
<th>INDC</th>
<th>FAVLC</th>
<th>VRL</th>
<th>Scheme presented</th>
</tr>
</thead>
<tbody>
<tr>
<td>S3578</td>
<td>48.02</td>
<td>45.12</td>
<td>45.05</td>
<td>52.15</td>
<td>52.14</td>
<td>58.46</td>
</tr>
<tr>
<td>S9234</td>
<td>43.59</td>
<td>42.79</td>
<td>47.08</td>
<td>45.82</td>
<td>50.17</td>
<td>58.86</td>
</tr>
<tr>
<td>S13207</td>
<td>81.30</td>
<td>80.43</td>
<td>85.27</td>
<td>81.58</td>
<td>83.29</td>
<td>84.87</td>
</tr>
<tr>
<td>S15850</td>
<td>66.22</td>
<td>65.13</td>
<td>67.24</td>
<td>67.70</td>
<td>69.78</td>
<td>70.70</td>
</tr>
<tr>
<td>S38417</td>
<td>43.26</td>
<td>56.52</td>
<td>47.61</td>
<td>43.06</td>
<td>62.84</td>
<td>50.46</td>
</tr>
<tr>
<td>S38584</td>
<td>60.91</td>
<td>60.57</td>
<td>66.02</td>
<td>72.29</td>
<td>65.42</td>
<td>71.38</td>
</tr>
<tr>
<td>Avg.</td>
<td>45.33</td>
<td>58.43</td>
<td>59.71</td>
<td>60.43</td>
<td>63.94</td>
<td>65.79</td>
</tr>
</tbody>
</table>
Table 5: Comparison with other algorithms (the hard fault set).

<table>
<thead>
<tr>
<th>Circuit</th>
<th>Size of TD bit</th>
<th>FDR</th>
<th>AR</th>
<th>ARL</th>
<th>Scheme presented</th>
</tr>
</thead>
<tbody>
<tr>
<td>S5738</td>
<td>5992</td>
<td>1110</td>
<td>1160</td>
<td>1091</td>
<td>710</td>
</tr>
<tr>
<td>S9234</td>
<td>7311</td>
<td>22038</td>
<td>21222</td>
<td>20429</td>
<td>11982</td>
</tr>
<tr>
<td>S13207</td>
<td>2205</td>
<td>14080</td>
<td>15534</td>
<td>15061</td>
<td>7500</td>
</tr>
<tr>
<td>S15850</td>
<td>163748</td>
<td>14370</td>
<td>13286</td>
<td>12861</td>
<td>15183</td>
</tr>
<tr>
<td>S38417</td>
<td>2201472</td>
<td>103760</td>
<td>94682</td>
<td>90397</td>
<td>103779</td>
</tr>
<tr>
<td>S38584</td>
<td>456768</td>
<td>13840</td>
<td>15696</td>
<td>15083</td>
<td>12551</td>
</tr>
</tbody>
</table>

By analyzing the experimental data, the relationship between the compression gain $\beta$ and the probability of not-care bit $\rho$ can be concluded as shown in Figure 3.

From Figure 3, as can be seen, the proposed compression scheme can obtain better compression effect in a higher probability of do not-care bit.

6. Conclusion

INS coding is presented, which uses the floating-point number unfolded by irrational number to store the test set. It is creative. Using the successive approximation method can accelerate the convergence speed in the process of searching the irrational number. The experimental results show that the compression effect of the scheme is better. It provides a new choice to solve the problem of test data compression.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors sincerely thank the anonymous reviewers for their valuable and constructive comments. This research is supported in part by the Natural Science Foundation of China (no. 61306046) and Scientific Research Foundation of High Institutions in Anhui Province (no. KJ2012B082).

References


