A Test Data Compression Method Based on K-L Transform

Yali Cheng

The College of Information Science and Engineering
Hunan University
Changsha, China
e-mail: cylhnu@163.com

Abstract—A test data compression method based on K-L transform is presented in this paper. The original test vector is divided into two parts by using K-L transform: the reference vector and the error vector. The reference vector is implemented approximately via Hadamard Matrix that has a simple hardware structure, and is integrated inside the circuit under test. The error vector is generated by using XOR between the reference vector and the original vector. In this paper, it encodes with the error vector, but without the original vector. In the procedure of synthesis, the reference vector generated by the hardware and the error vector produced by the decompression synthesize the original test vector. The experimental results show that this scheme can achieve an average compression ratio of 77.90%, improve about 15.36%~20.68% comparing with others.

Keywords-K-L transform; test data compression; the reference vector; the error vector; Hadamard Matrix

I. Introduction

With the rapid development of IC fabrication processing, there is an increasing number of IP cores on a single SOC(System-of-a-chip) and a growing complexity of the chip. When testing for the chip, a large number of test data is required to guarantee high fault coverage. However, the enormous data volume requires an increasing performance of the automatic tester (ATE) and longer test application time. Therefore, how to reduce the volume of test data has become a great challenge of IC testing.

Test data compression schemes can not only significantly decrease the amount of test data stored on the tester, but also effectively reduce the test application time for a given bandwidth. Test data compression schemes fall broadly into three categories [1]. Firstly, linear-decompression-based schemes decompress the data using only linear operations (that is LFSRs and XOR networks), however, the linear operations should have resolutions (that the test set can be encoded) if the number of free variables is 20 more than the number of specified bits [2]. Secondly, broadcast-scanbased schemes rely on broadcasting the same values of external scan chains to a large amount of internal scan chains, but the disadvantage of such schemes is the dependency between the internal scan chains and the external scan chains. Thirdly, code-based schemes use data compression codes to encode test cubes. For instance, Chandra [4] proposed FDR code to encode the runs of 0s in the test cubes, and Ei-Maleh [5] exhibited EFDR code to encode the runs of 0s and 1s simultaneously, and Liang [6] showed Alt-FDR code to

Jishun Kuang

The College of Information Science and Engineering

Hunan University

Changsha, China

e-mail: jshkuang@hotmail.com

encode the runs of 0s and 1s alternately. However, the existence of entropy limits the test compression [7].

Spectral analysis has been introduced into varied fields recently, such as test generation, test data compression. It deals with test vectors based on the characteristics of test vectors analyzing in the transform domain. Susskind [8] verified the coefficients of the circuit outputs by using Walsh spectrum to test a digital circuit for stuck-at faults. Hsiao and Seth [9] further expanded his work to compact the test set by selecting the Rademacher-Walsh coefficient of test set as the signature of output response. Khan and Bushnell [10][11] used spectral analysis to design the hardware architecture of the output response.

There are less literatures on spectral analysis in the field of test data compression. In this paper, we combine the spectral analysis and test data compression to propose a test compression mothed based on K-L transform. The main work includes: the original test vector is divided into the reference vector and the error vector by using K-L transform; the reference vector is represented and generated by K-L transform matrix, and the error vector is encoded; the K-L transform matrix is hard to be produced and the hardware cost is high, hence it is replaced with Hadamard matrix that has a simple structure; in the procedure of synthesis, the reference vector generated by the hardware and the error vector produced by the decompression synthesize the original test vector. The experimental results show that this scheme can effectively improve the test compression ratio.

II. K-L TRANSFORM

A. Transform coding

Transform coding realizes data modeling expression via mapping transformation, and its general model is shown in Fig.1. The mapping transformation transforms the original data from one domain to another domain, such as the signal is transformed from time domain to frequency domain, then encodes the data in another domain [12]. What is important is that the mapping transformation should produce a series of effective coefficients. When encoded, the total bits required to these coefficients are less than the bits required to the original data, thus the data can be efficiently compressed.



Figure 1. General model of transform coding

Orthogonal transformation method is used frequently, even though there are a large number of mapping transformation schemes. Fourier transform converts a function from time domain to frequency domain through complex domain orthogonal transform, thus simplifies the problem. Common transform coding is Krhunen-Leove transform (K-L transform), Hadamard transform, DCT transform and so on.

B. K-L transform

K-L transform is based on statistical characteristics and featuring good de-correlation. It has been recognized as the best transform in the means square error (MSE) and an important technique in test data compression as well.

The orthogonal matrix formed by the normalized orthogonal eigenvector of the covariance matrix of the test set *T* is called the K-L transform matrix *A*. The K-L transform is shown in the following formula (1), and the inverse transform is shown in the following formula (2).

$$Y=AT$$
 (1)

$$T=A'Y$$
 (2)

Here, A' means the transpose of the K-L transform matrix A.

Although K-L transform has excellent performance on data compression, the amount of calculation required to calculate the covariance matrix and eigenvectors is large. So K-L transform is limited in practical use. The approach presented in this paper analyzes the characteristics of the original test set by using K-L transform, then discusses its theory compression ratio.

C. Hadamard Transform

The Hadamard transform was proposed to analyze the frequency characteristics for binary digital signals, which is similar to the Fourier transform for the analog signal [13]. While the Fourier transform uses sine functions as their basis, the Hadamard transform utilizes a series of orthogonal functions called Walsh functions.

The Hadamard transform and its inverse transform can be expressed:

$$Y=HT$$
 (3)

$$T=HY$$
 (4)

Here, H represents the Hadamard matrix, and its inverse matrix is itself.

Walsh function contains only two values +1 and -1, corresponding to two states of the digital logic. Any binary bit-stream can be uniquely represented as a linear combination of the orthogonal Walsh functions, which analogous to the analog domain where any continuous signal can be uniquely represented as a linear combination of the sine and cosine functions. Consequently, we actually study the frequency characteristics of the digital waveform by analyzing the binary bit-stream using Walsh functions.

Hadamard transform matrix containing Walsh functions can be defined in two ways: the binary representation and the recursive representation. In this paper, Hadamard matrix is represented with the recursive definition:

$$H(n) = \begin{bmatrix} H(n-1) & H(n-1) \\ H(n-1) & -H(n-1) \end{bmatrix}$$

H(0)=1, and 2^n is the order of the nth Hadamard matrix H(n),. For example, when n=2, H(0) can be shown as:

III. THE ALGORITHM

A. The main algorithm

Code-based schemes use data compression codes to encode the test cubes. This involves in the don't care bits X filling. In this paper, the original test vector is divided into the reference vector and the error vector by using K-L The reference vector is implemented transform. approximately by using Hadamard Matrix that has a simple hardware structure, with integrated inside the circuit under test. The error vector is generated by using XOR between the reference vector and the original vector. In this paper it encodes with the error vector, but without the original vector. In the synthesis, the reference vector generated by the hardware and the error vector produced by the decompression make up the original test vector.

The rows of the original test set refer to the number of the test vector, and the columns refer to the bits of the test vector. For instance, the original test set of the benchmark circuit s5378 has 111 test vectors that each has 214 bits, which means this test set has 111 rows and 214 columns. The test set consists of the specific bits 0, 1 and the don't care bits X. When compressing, the specific bits usually cannot be changed, otherwise it will cause the loss of the fault coverage. For the stuck-at faults, the bit sequence among each test vector cannot be changed in the process of testing. If not, it may reduce the fault coverage. But the sequence among different test vectors can vary during the testing. It is shown that the correlation among the bits in each test vector is larger than the correlation among different test vectors. Therefore, the correlation among the columns (i.e. bits among each vector) is larger than the correlation among the rows (i.e. different test vectors) in the original test set.

Since the good de-correlation of K-L transform, the results that K-L transform performs on the columns of the original test vector should be better than that on the rows theoretically. Consequently, the method proposed in this paper analyzes the original test set by using K-L transform on the columns and rows respectively. The procedure of K-L transform on the columns of the original test vector is considered as vertical K-L transform, and the procedure of K-L transform on the rows as horizontal K-L transform. Before utilizing K-L transform, it's necessary to calculate the K-L transform matrix of the original test set. The calculation is described in the following example of a vertical K-L transform. Assuming that the size of the original test set T is calculated as following:

 $C_T = E\{T - E(T) [T - E(T)]^T\}$

E(T) is the mean value of the original test set T. We compute the eigenvalues λ_k of the covariance matrix in order and the

eigenvectors u_k which satisfy the following conditions:

$$u_l^T u_k = \begin{cases} 1, l = k \\ 0, l \neq k \end{cases}$$

Thus the K-L transform matrix A is an orthogonal matrix composed by eigenvectors, and its size is m*m.

The procedure of the vertical K-L transform is as follows. Firstly, for the first column of the test set, K-L transform is applied to extract the K-L coefficient matrix Y, and Y whose size is m*1 shows the energy distribution of the first column in the transform domain. For instance, the first number of Y represents the first column's energy based on the first row of the K-L transform matrix, and the second number of Y reports the first column's energy based on the second row of the K-L transform matrix, and so on. Secondly, the row matching the original test vector best is selected from the K-L transform matrix as its reference vector. Therefore, the row where the maximum value from the coefficient matrix Y in is obtained, and the corresponding row from the K-L transform matrix is transposed as the reference vector of the first column. Finally, the remaining columns in the original test set is coped with the same process, thus obtaining a reference matrix B. Note that the matrix B is generated by selecting the corresponding rows from the K-L transform matrix A based on the coefficient matrix Y.

The error matrix D is the result of XOR between the original test set T and the reference matrix B. In this approach, the XOR has slight difference with the general XOR. When comparing the specific bit in the original test set T with the corresponding bit in the reference matrix B, the result is 0 if the two are same, or it is 1. When comparing the don't care bit X in the original test set T with the corresponding bit in the reference matrix B, the result is the don't care bit X whether the two are same or not. Obviously, the position where the don't care bits are in the original test set is same with the position in the error matrix. In this paper, the error matrix D is encoded to compress, but not the original test set T.

B. Preprocess and the discretization of the K-L transform matrix

The original test set of the ISCAS'89 benchmark circuits consists of the specific bits 1,0 and the don't care bits X, while the K-L transform matrix A is a real matrix based on the covariance matrix. Thus the original test set and the K-L transform matrix should be preprocessed before transforming. Since the don't care bit X has no contribution to the energy distribution of the whole test set, the 0s in the original test set are set to the -1s, and the Xs are set to the 0s, and the 1s remain. The Xs are usually filled with the 0s or the 1s during compressing, in another word, the filled test set just consists of 0s and 1s. In this paper, the K-L transform matrix with the real style is discretized to a binary matrix by setting a threshold. The process is called the discretization of the K-L transform matrix.

To select more similar reference vector to the original test vector, the K-L transform matrix after discretized should be more similar to the original test set. According to the characteristic of the original test set that the ratio of the 0s and 1s is close to 1, suppose when the ratio of the 0s and 1s

in the K-L transform matrix is closer to 1, the compression will be better. Thus the proposed method chooses several different thresholds to show the relationship between the compress ratio and the thresholds.

In this paper, the mean value of the original test set is viewed as a threshold for the K-L transform matrix, called the mean threshold. For the vertical K-L transform, an example of this threshold is to show the algorithm idea presented in the paper. Assuming that the original test set T as following, T only contains the specific bits for convenience:

$$T = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Firstly, set the 0s in the original test set T to the -1s as following:

Secondly, calculate the covariance matrix C and the K-L transform matrix A:

$$C = \begin{bmatrix} 0.8889 & -0.6667 & -0.6667 & 0.4444 \\ -0.6667 & 1 & 0.3333 & -0.3333 \\ -0.6667 & 0.3333 & 1 & -0.3333 \\ 0.4444 & -0.3333 & -0.3333 & 0.5556 \\ \end{bmatrix}$$

$$A = \begin{bmatrix} -0.8077 & -0.3771 & -0.3771 & 0.2514 \\ 0 & 0.3015 & 0.3015 & 0.9045 \\ 0 & 0.7071 & -0.7071 & 0 \\ -0.5897 & 0.5166 & 0.5166 & -0.3444 \end{bmatrix}$$

Thirdly, gain the mean value of the original test set whose value is 0.2500 as the mean threshold, then set the values more than the threshold to 1 and the values less than the threshold to -1. So the discretized K-L transform matrix *A* is shown as following:

Fourthly, transform the first column of the original test set to obtain the coefficient matrix *Y*:

$$Y = \begin{bmatrix} -2 \\ 2 \\ 2 \\ 4 \end{bmatrix}$$

Fifthly, select the row from the K-L transform matrix A as the reference vector according to the row where the maximum value of the coefficient matrix Y in, i.e. the row 4. Take the same measures to the remaining columns in the original test set. Thus the reference matrix B is represented as:

Finally, the error matrix D is the result of XOR between the original test set T and the reference matrix B, shown as following:

As a result, the original test set is divided into two parts by using K-L transform: the reference matrix B and the error matrix D. As it can be seen, the high comparability between the original test vector and the reference vector leads to the growing number of 0s in the error vector. Consequently, in theory, the compression ratio obtained by encoding the error vector should be higher than that encoding the original test vector.

C. Approximate implementation

Based on the coefficients, the reference vector is selected from the K-L transform matrix. Therefore, what important to implement the reference vector is to implement the K-L transform matrix. However, K-L transform involves in the calculation of the covariance matrix and eigenvectors which is complex and large, resulting in implementing the hardware architecture for K-L transform difficultly. In this approach, Hadamard matrix that has a simple structure is used to implement the K-L transform matrix approximately.

The procedure implemented by Hadamard matrix is similar to that by K-L transform, and the difference for Hadamard is that the transform matrix is Hadamard-Walsh matrix, so it's not repeated here. But it may cause the compression ratio drop.

IV. SYNTHESIS

As explained earlier, the original test vector is broken down to two parts to compress. The test vector should be decompressed before testing the circuit under test. The error vector is decompressed based on the coding scheme it utilized, and the reference vector is generated by the Hadamard-Walsh Generator [14] inside the circuit under test. Comparing with the original test vector, the error vector contains more 0s. Since the FDR scheme encodes the runs of 0s in the test cube, the FDR scheme is chose to encode the error vector in this paper. The decompression structure of the error vector is the same as the FDR decompression structure[4], so it is not repeated here. In this paper, the vector synthesized by the error vector and the reference vector is called the synthesis vector.

V. EXPERIMENTAL ANALYSIS

Experiments were carried on the ISCA'89 benchmark circuits for investigating the efficiency of the proposed method for data compression based on the algorithm of K-L transform. Here only results for the larger benchmark circuits are reported. The experimental results in Table 1 compare the compression ratio between the vertical K-L transform and the horizontal K-L transform based on the proposed method. The length of test vectors and the number of test vectors are represented in columns 2 to 3. This experiment used the mean threshold and FDR code. The compression ratio of the FDR code without using transform coding [4], the vertical K-L transform, and the horizontal K-L transform

are listed in the last 3 columns. It can be observed that the compression ratio of the proposed method is higher than that of the FDR code. Also note that the average compression ratio of the vertical K-L transform is better than the horizontal K-L transform, improves about 6.10%. This is due to the fact that the correlation among the columns (i.e. bits among each vector) in the original test set is larger than the correlation among the rows (i.e. different test vectors), and the K-L transform has a good de-correlation. Therefore, the vertical K-L transform is chose to perform in the next experiments.

TABLE I. THE COMPRESSION RATIO USING K-L TRANSFORM

	Length	Number	FDR	Compression ratio of the proposed method(%)	
Circuit	of vectors	of vectors	[4]	Vertical K-L transform	Horizontal K-L transform
S5378	214	111	48.02	71.41	64.95
S9234	247	159	43.59	69.92	62.22
S13207	700	236	81.30	91.91	90.46
S15850	611	126	66.22	82.98	78.67
S38417	1664	99	43.26	74.83	58.65
S38584	1464	136	60.91	76.34	75.84
The ave	The average compression ratio			7790	71.80

The results in Table 2 show that the proposed method compares with other schemes [7]. Here the length of test vectors and the number of test vectors are omitted. The columns 2 to 4 report the compression ratio of the FDR code[4], the EFDR code[5] and the Alt-FDR code[15] respectively. And the proposed vertical K-L transform is repeated in the last column. Please note that in all cases, our method performs better than other schemes, the average compression ratio achieves 77.90%. As it can be seen from the results, our method improves the compression ratio between about 20.68% and 15.36%. The experimental results show that the vertical K-L transform technique has an effectivity and feasibility on test data compression.

TABLE II. THE COMPRESSION RATIO AMONG THE PROPOSED METHOD AND OTHERS

	The compression ratio(%)				
Circuit	FDR[4]	EFDR[5]	Alt-FDR[15]	The proposed vertical K-L transform	
S5378	48.02	51.93	-NA	71.41	
S9234	43.59	45.89	44.96	69.92	
S13207	81.30	81.85	80.23	91.91	
S15850	66.22	67.99	65.83	82.98	
S38417	43.26	60.57	60.55	74.83	
S38584	60.91	62.91	61.13	76.34	
The average	57.22	61.86	62.54	77.90	

To analyze the influence of different thresholds on the test compression, the results of the proposed method under different thresholds are provided in Table 3. In this experiment, we cope with four thresholds. The threshold 1 refers to the means of the original test set called the means threshold, and the threshold 2 refers to the means of the K-L transform matrix called the average threshold. The threshold

3 describes the 3/4 threshold close to the maximum within the region between the maximum and the minimum of the K-L transform matrix. The threshold 4 depicts the 1/4 threshold close to the minimum within the region between the maximum and the minimum of the K-L transform matrix. So the threshold 2 is located in the middle position between the threshold 3 and the threshold 4. The columns 2 to 4 list the compression ratio under the threshold 1, the threshold 2, the threshold 3 and the threshold 4. As shown by our experimental results, the compression ratio under the threshold 1 and the threshold 2 is higher than that under the threshold 3 and the threshold 4. The main explanation lies in a fact that under the threshold 1 and the threshold 2, the ratio of 0 and 1 in the K-L transform matrix is closer to 1.as mentioned before, thus provides more potential for compression. Furthermore, it is also observed that the results under the threshold 1 are better than that under the threshold 2. The K-L transform matrix is based on the characteristic of the original test set, and the mean value of the original test set is more suitable for discretizing the K-L transform matrix than the mean value of the K-L transform matrix. The experimental results in Table 3 show that thresholds have influence on the data compression based on the proposed method, and the mean threshold performs best. Thus the mean threshold is used to discretize the K-L transform matrix.

TABLE III. THE COMPRESSION RATIO UNDER DIFFERENT THRESHOLDS

Circuit	The compression ratio under different thresholds(%)					
Circuit	Threshold 1	Threshold 2	Threshold 3	Threshold 4		
S5378	71.41	70.12	65.36	64.58		
S9234	69.92	69.78	61.42	61.00		
S13207	91.91	91.70	87.88	87.90		
S15850	82.98	82.62	76.81	76.86		
S38417	74.83	74.64	71.70	71.67		
S38584	76.34	76.32	69.08	69.09		
The average	77.90	77.53	72.04	71.85		

Due to the complex implementation of K-L transform, the simple Hadamard matrix is applied to implement the K-L transform matrix approximately in this approach. The comparison of the K-L theoretical compression ratio and the Hadamard implementation compression ratio is reported in Table 4. The compression ratios obtained by the FDR

TABLE IV. THE RESULTS COMPARING THE K-L TRANSFORM AND THE HADAMARD IMPLEMENTATION

	The compression ratio(%)				
Circuit	FDR[4]	The vertical K-L transform	The Hadamard implementation		
S5378	48.02	71.41	64.65		
S9234	43.59	69.92	61.70		
S13207	81.30	91.91	86.55		
S15850	66.22	82.98	75.77		
S38417	43.26	74.83	69.95		
S38584	60.91	76.34	72.63		
The average	57.22	77.90	71.88		

code, the vertical K-L transform and the Hadamard matrix are listed in the columns 2 to 4. From the Table 4, it can be found that the compression ratio drops after using the Hadamard matrix to implement the K-L transform matrix, which mainly due to the K-L transform matrix implemented by the Hadamard matrix has error with the original transform matrix. However, the Hadamard scheme still achieves 71.88%, improves 14.66% comparing to the FDR codes.

VI. CONCLUSION

A test data compression method based on the K-L transform has been presented in this paper. This scheme improves the compression ratio effectively by dividing the original test vector into two parts. The error vector can be encoded more efficient than the original vector. But the reference vector is implemented approximately by the Hadamard matrix. It reduce the hardware cost by replacing the complex K-L transform matrix with the simple Hadamard matrix, and the loss of the compression ratio is within acceptable limits.

- N. Tauba, "Survey of test vector compression techniques," IEEE Transaction Design & Test of Computers, pp. 294

 –303,2006.
- [2] B. Koenemann, "LFSR-Coded Test Patterns for Scan Designs," Proc. European Test Conf. (ETC 91), VDE Verlag, 1991, pp. 237-242.
- [3] A.H.El-Maleh. efficient test compression technique based on block merging[J]. IET compter& Digital Techniques, 2007, pp.327-335.
- [4] Chandra A,Chakrabarty K. Frequency directed run-length(FDR) codes with application to system on-a-chip test data compression[J].VLSI Test Symposium,2001,pp.42-47.
- [5] EI-Maleh A,Al-Abaji R.Extended frequency directed run-length code with improved application to system data compression[J]. International Conference on Electronic Circuits and Systems,Mar.2002, pp.449-452.
- [6] Liang H G, Jiang C Y. Efficient test data compression and decompression based on alternation and run-length codes[J]. Chinese Journal of Computers, Apr. 2004, pp. 548-554.
- [7] Usha S.Mehta, Kankar S.Dasgupta, and Niranan M.Devashrayee, "Run-Lenth-Based Test Data Compression Techniques: How Far from Entropy and Power Bounds? – A Survey"[J]. VLSI Design,Jan.2010, pp.1-9.
- [8] A. K. Susskind, "Testing by Verifying Walsh Coefficients," IEEE Trans. Computers, vol. 32, no. 2, pp.198-201, Feb. 1983.
- [9] T.-C. Hsiao and S. C. Seth, "An Analysis of the Use of Rademacher-Walsh Spectrum in Compact Testing," IEEE Trans. Computers, vol. 33, pp. 934-938, Oct.1984.
- [10] O. I. Khan and M. L. Bushnell, "Aliasing Analysis of Spectral Statistical Response Compaction Techniques," in Proc. 19th International Conf. on VLSI Design, Jan. 2006, pp. 801-806.
- [11] O. I. Khan, M. L. Bushnell, S. K. Devanathan, and V. D. Agrawal, "SPARTAN: A Spectral and Information Theoretic Approach to Partial Scan," in Proc. International Test Conf., Oct.2007. Paper 21.1.
- [12] Lenan Wu, Tata Compression, Electronic Industry Press, Oct.2005.
- [13] F.R.Ohnsorg, "Binary Fourier representation," in Spectrum Analysis Techniques Symposium, Sept.1966.
- [14] C. K. Yuen, "New Walsh-Function Generator," Electronics Letters, vol. 7, p. 605, 1971.
- [15] A. Chandra and K. Chakrabarty, "Reduction of SOC test data volume, scan power and testing time using alternating runlength codes," in Proceedings of the 39th Design Automation Conference (DAC '02), pp. 673–678, New Orleans, La, USA, June 2002.