

一個低成本的二維雙正交離散小波之超大型積體電路架構設計

游竹¹ 陳少傑²

1. 國立宜蘭技術學院電子工程系助理教授
2. 國立台灣大學電機工程系教授

摘要

類似於短時間富立葉轉換，小波轉換提供了另一種信號處理的方法，特別的是它適合於空間及頻譜聚集性的分析。在本論文中，我們提出一個低成本的二維雙正交離散小波轉換之超大型積體電路架構設計，藉由有效的安排計算排程及運用雙正交濾波器的對稱特性，因而得到硬體成本的最小化。針對計算一個含 L 長度的 $N \times N$ 二維影像而言，我們所提出的架構需花費近 N^2 時脈週期，另外要求約 $2NL - 2N$ 的存儲單元、如為偶長度的濾波器則需 $15L/8$ 個乘法器、如為奇長度的濾波器則需 $15L/2 + 1/4$ 個乘法器，以及 $4(L-1)$ 個加法器等計算單元硬體。

關鍵詞：離散小波轉換、超大型積體電路

Design of a Low-Cost VLSI Architecture for 2-D Biorthogonal Discrete Wavelet Transforms

Chu Yu¹ and Sao-Jie Chen²

1. Assistant Professor, Department of Electronic Engineering, National Ilan Institute of Technology
2. Professor, Department of Electrical Engineering, National Taiwan University

Abstract

The wavelet transform, similar to the Short-Time Fourier Transform, furnishes an alternative approach to signal processing, especially suitable for the analysis of spatial and spectral locality. However, the 2-D wavelet transform needs the same enormous computation as their counterpart Fourier transforms. In this paper, we present a low-cost VLSI architecture for 2-D biorthogonal discrete wavelet transforms. Both the efficient arrangement of a computation schedule and the symmetric property of biorthogonal filters used in this architecture contribute to the hardware cost minimization. For the computation of an $N \times N$ 2-D image with a filter length L , this architecture spends near N^2 clock cycles, and requires about $2NL - 2N$ storage unit, $15L/8$ multipliers for even-length filters or $15\lfloor L/2+1 \rfloor/4$ multipliers for odd-length filters, as well as $4(L-1)$ adders.

Key Words : VLSI architecture, discrete wavelet transforms

I. Introduction

The wavelet transform (WT), similar to the Short-Time Fourier Transform (STFT), furnishes an alternative approach to signal processing, especially suitable for the analysis of spatial and spectral locality [1]-[5]. The WT differs with the STFT in their time-frequency representations. By processing data with different window widths at variable scales, it thus overcomes the limitation of STFT whose time-frequency resolutions are fixed over the entire time-frequency plane. However, since 2-D DWT's need the same enormous computation as their counterpart Fourier transforms, in order to meet the requirements of fast computation in many real-time applications, dedicated hardware implementations are necessary.

Several VLSI architectures based on the separable approach for 2-D discrete wavelet transforms have been proposed [6]-[18]. For instance, Lewis and Knowles [6] firstly proposed a multiplierless 2D architecture for the 4filter-length Daubechies wavelet transform. Vishwanath et al. [7] presented a 2D systolic-parallel DWT architecture that uses a combination of systolic and parallel filters. Chakrabarti et al. [8] devised many efficient architectures, such as parallel filters (in 1-D and 2-D), SIMD linear array, and SIMD multigrid architectures, which utilize parallel structures to implement the existing pyramid [1] and "à trous" [9] algorithms for the DWT and the continuous WT, respectively. Following the above works, Denk and Parhi [10] presented an implementation of the 2-D DWT based on lapped block processing techniques. A parallel-pipelined VLSI array architecture for the 2-D DWT has also been proposed by Chuang and Chen [11]. Rumian [12] described a practical implementation of the 2-D DWT decomposition based on a pair of 1-D QMF FIR filters and two FPGA circuits. Recently, Chen and Bayoumi [13] presented a systematic synthesis approach for scalable systolic array architecture for the 2-D DWT based on the data dependence analysis and linear index space transformation. Later, to improve the foregoing architecture, Limqeco and Bayoumi [14] again proposed an efficient semi-systolic array architecture for 2-D discrete wavelet transforms. This architecture hold the same amount of computational unit as the systolic array architecture, but its communication routing has lower complexity because the data transfer between processing elements is limited to its neighbors and data to be stored within the processor's local memory. More recently, Sheu et al. [15] proposed a lower hardware cost and less memory VLSI architecture for the separable approach of 2-D DWT, which provides a low on-chip memory needed. Paek et al. [16]-[17] presented a two-stage semi-recursive VLSI architecture for the 2-D DWT.

All the architectures described above have to consider the trade-off between the cost of their computation units and the time spent for computing 2-D DWT's. To alleviate this problem, we present a VLSI architecture for 2-D DWT's, which owns low-cost computation units and fast computation time. We mainly apply an efficient computation schedule and take advantage of the symmetric property of a biorthogonal filter bank in our proposed architecture to reduce its hardware cost needed. The purpose of our study is aimed at meeting the requirement of low-cost hardware implementation for the computation unit, so that we can use FPGA or single-chip VLSI to implement the hardware of 2-D DWT's.

II. 2-D Discrete Wavelet Transforms

In the separable approach, the 2D wavelet transforms can use cascading 1D wavelet transforms to perform row/column-major transform and then column/row-major transform. Assume that the input source is an image, then the image can be first decomposed into a pair of highpass and lowpass subimages by applying a 1-D wavelet transform in the row dimension. Then, the highpass and lowpass subimages with downsampled outputs can be further decomposed in the column dimension into a new pair of coarser lowpass and highpass subimages, respectively, by using another 1-D wavelet transform. Therefore, the four resultant sub-images in the output contain information of different significance, and the size of each sub-image is reduced to a quarter of the original image after the decomposition processing. In other words, we obtain now four outputs HH , HG , GG , and GH , where the sub-image HG is the output data via the lowpass row filter H and the highpass column filter G , and the other HH , GG , and GH sub-images are generated similarly. Further analysis of the image will decompose the subimage HH into four new sub-images, and so on, until the coarsest level of sub-images are achieved. One's decomposition with two resolution levels is shown in Fig 1.

As described above, the decomposition results for an image using 2-D DWT are shown in Fig 2. The left of the figure corresponds to its individual scale after three levels of 2-D DWT decomposition; and the right of the figure is a practical example for the Lena image. Through such decomposition, the highest-frequency components of the entire image are concentrated in the $GG1$ that represents the finest subimage in the image. In contrast with the $GG1$, the lowest-frequency

components of the entire image are concentrated in the $HH3$ that shows the coarsest subimage in the image. Thus, the image is classified into various resolution levels corresponding to its different scales. Due to these properties, the 2-D discrete wavelet transform is frequently used in many applications for image processing, such as image compression, image recognition, edge detection, image de-noising, and so on.

However, since the amount of the 2-D DWT computation is similar to other classical 2-D transforms, such as the FFT or DCT, the computational complexity of the 2-D DWT is also too expensive. To satisfy the requirements in real-time image processing such as on-line videophone and videoconference, a dedicated hardware implementation is necessary. In the following sections, we will propose a low-cost VLSI architecture for the 2-D DWT computations.

III. 2-D DWT Architectures

1. Proposed VLSI architecture

The proposed VLSI architecture for 2-D biorthogonal DWT is shown in Fig 3. The architecture consists mainly of four parallel filters and one storage-unit. Here, the four filters are responsible for performing the row and column filterings, respectively. The storage unit is used for storing the row filtering outputs, in which data will be transposed to serve as the input of the column filter. By using an efficient computation schedule (as described in the next subsection) to compute DWT's on-the-fly, the storage unit between resolution levels can be eliminated. Moreover, when we use a delay-line scheme to implement the storage unit, the delay-line used in the last filter-tap can also be eliminated because these input samples were already available at the input terminal and we do not need to store them in a buffer. As a result, the size of the storage unit required in our proposed architecture is $2NL-2N = 2N(L-1)$ for storing both highpass and lowpass filterings, where N is the row/column size of an image and L is the filter length.

2. Computation Schedule

A computation-schedule table illustrating the three-level 2-D DWT computations for an $N \times N$ image is shown in Table 1. In the table, H_i and V_i represent the horizontal and vertical filtering computations at the resolution level i , respectively. For simplicity, only the V_i computations of a single vertical filter (may be $VF1$ or $VF2$) are shown in this table. This schedule table is generated from the data dependence analysis of the separable 2D DWT computations. A horizontal filtering computation is shown in the upper side of a grid in the table and a vertical filtering computation is shown in the lower side. An upper- or a lower-sided blank grid represents no filtering computation there. Note that each of the last computation points of all levels will be wrapped to its next row except for $H1$. For instance, the last computation points of each row of $H2$ are wrapped to grid (2, 1), for $l = 2, 4, \dots, N$; and the last vertical computation points of each column of $V1$ appear at grid (0, 1), for $l = 2, 4, \dots, N$. Therefore, all the resolution levels of computations shown in this schedule table are interleaved with the exact spacing of the first-level.

The idea of this computation schedule is that the generation sample is computed if all its input samples are available and the computation unit is idle currently. For example, the horizontal computation $H2$ at grid (1, 6) is generated because the vertical computation $V1$'s at grids (1, 2) and (1, 4) are available and the computation unit of the horizontal filtering ($HF1$) is idle.

From the above table, the total computation time for computing the three levels of $N \times N$ 2-D DWT is $N^2 + 6 \approx N^2$ clock cycles. Thus, this computation schedule is efficient because only a small amount of extra DWT computations is needed, which is far less than the existing algorithm [7]. In addition, as mentioned earlier, when using our proposed computation schedule, not only the storage buffer between resolution levels for the next level of DWT computation, but also the delay line used in the last filter tap can be eliminated. Therefore, the size of the storage units used in our proposed architecture is further reduced.

3. Filter Structure

To further reduce the hardware cost of the computation unit, we design an efficient filter structure for the biorthogonal filter bank, as shown in Fig 4. Here, D is a delay element. Symbols g and h are highpass and lowpass coefficients, respectively. The biorthogonal filter bank has the symmetric property, e.g. coefficients $g_0 = g_4$, $g_1 = g_3$, $h_0 = h_4$, and $h_1 = h_3$

for a five-tap filter as shown in Fig. 4, thus we can reduce about half of the multipliers needed. Based on this filter structure, our proposed architecture can further reduce the hardware cost of its computation units, such that it is suited for FPGA or single-chip VLSI implementation.

IV. Performance Evaluation

The performance evaluation of various architectures is summarized in Table 2, where N denotes the number of row/column input samples (8-bit data width) and L is the filter length. Period is defined as the minimum time between the first input of one computation and the first input of the next computation. Assuming that multipliers of 16×16-bit precision are used for all the other architectures in the table. For a better implementation, we optimize the precision requirements of the multipliers in our proposed architecture. That is, the multipliers used in the filter $HF1$ need only a precision of 8×16 bits, while the ones used in the remaining filters ($HF2$, $VF1$ and $VF2$) need a precision of 16×16 bits. The reason is that $HF1$ performs only the computation of input samples, thus it requires lower precision than other filters. In total, the number of multipliers required in our architecture becomes $\frac{L}{2} + \frac{L}{2} + \frac{L}{2} + \frac{3L}{8} = \frac{15L}{8}$ for even-tap filters and $3\lfloor\frac{L}{2}+1\rfloor + \frac{3}{4}\lfloor\frac{L}{2}+1\rfloor = \frac{15}{4}\lfloor\frac{L}{2}+1\rfloor$ for odd-tap filters. Clearly, from Table 2, our proposed architecture requires a smaller number of computation units and computation time than other existing architectures.

V. Conclusions

A low-cost VLSI architecture for 2-D biorthogonal DWT was described in this paper. This architecture with its efficient computation schedule can reach about a period of N^2 . We also designed a low-cost filter structure and optimized the precision requirement of processing elements to reduce the hardware cost. Since the architecture has fast computation time and low-cost hardware, it is suited for single-chip VLSI implementation, especially for the rapid prototyping using FPGA. In addition, it can also be applied very well to many real-time video/image applications, such as MPEG-4 and JPEG 2000.

Acknowledgment

This research was supported by the National Science Council, Taiwan, ROC, under Grant NSC90-2215-E-197-001.

References

1. S. Mallat, (1989), "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 11, no. 7, pp. 674-693.
2. Rioul and M. Vetterli, (1991), "Wavelets and signal processing," *IEEE Signal Processing Magazine*, vol. 8, no.4, pp. 14-38.
3. Daubechies, (1992), Ten Lectures on Wavelets, vol. 61 of CBMS-NSF Regional Conferences Series in Applied Mathematics, SIAM, Philadelphia, PA.
4. M. Vetterli and J. Kovacevic, (1995), Wavelet and Subband Coding, Prentice Hall PTR, Englewood Cliffs.
5. G. Strang and T. Nguyen, (1996), Wavelets and Filter Banks, Wellesley-Cambridge Press.
6. A. S. Lewis and G. Knowles, (1991), "VLSI architecture for 2-D daubechies wavelet transform without multipliers," *Electron. Lett.*, vol. 27, no. 2, pp. 171-173.
7. M. Vishwanath, R. Owens, and M. J. Irwin, (1995), "VLSI architectures for the discrete wavelet transform," *IEEE Trans. Circuits and Systems II, Analog and Digital Signal Processing*, vol. 42, no. 5, pp. 305-316.
8. C. Chakrabarti and M. Vishwanath, (1995), "Efficient realizations of the discrete and continuous wavelet transforms: from single chip implementations to mappings on SIMD array computers," *IEEE Trans. Signal Processing*, vol. 43, no. 3, pp. 759-771.
9. M. Holschneider, R. Kronland-Martinet, J. Morlet, and Ph. Tchamitchian, (1989), "A real-time algorithm for signal

- analysis with the help of the wavelet transform,” in *Wavelets, Time-Frequency Methods and Phase Space*, J. M. Combes, A. Grossmann, and Ph. Tchamitchian, Eds. Berlin: Springer, IPTI, pp. 286-297.
10. T. C. Denk and K. K. Parhi, (1994), “Calculation of minimum number of registers in 2-D discrete wavelet transforms using lapped block processing,” in *Proc. IEEE Int’l Symp. on Circuits and Systems*, May, pp. 524-527.
 11. H. Y. H. Chuang and L. Chen, (1995), “VLSI architecture for fast 2D orthonormal wavelet transform,” *Journal of VLSI Signal Processing*, vol. 10, no. 3, pp.225-236.
 12. R. Rumian, (1994), “An architecture for real-time wavelet image decomposition,” in *Proc. IEEE Int. Symp. on Circuits and Systems*, London, England, pp. 73-76.
 13. J. Chen and M. A. Bayoumi, (1995), “A scalable systolic array architecture for 2-D discrete wavelet transforms,” in *Proc. IEEE Workshop on VLSI Signal Processing*, vol. III, pp. 303 –312.
 14. J. C. Limqueco and M. A. Bayoumi, (1998), “A VLSI architecture for separable 2-D discrete wavelet transform,” *Journal of VLSI Signal Processing*, vol. 18, no.2, pp.125-140.
 15. M. H. Sheu, M. D. Shieh, and S. W. Liu, (1998), “A VLSI architecture design with lower hardware cost and less memory for separable 2-D discrete wavelet transform,” in *Proc. IEEE Int. Symp. on Circuits and Systems*, pp. 457-460.
 16. S. K. Paek, H. K. Jeon, and L. S. Kim, (1998), “Semi-recursive VLSI architecture for two-dimensional discrete wavelet transform,” in *Proc. IEEE Int. Symp. on Circuits and Systems*, pp. 469-472.
 17. S. K. Paek, H. K. Jeon, and L. S. Kim, (1998), “2D DWT VLSI architecture for wavelet image processing,” *Electron. Lett.*, vol. 34, no. 6, pp. 537-538.
 18. C. Yu and S. J. Chen, (1999), “Design of an efficient VLSI architecture for 2-D discrete wavelet transforms” *IEEE Trans. on Consumer Electronic*, vol. 45, no. 1, pp. 135-140.

91 年 08 月 02 日投稿

91 年 08 月 20 日接受

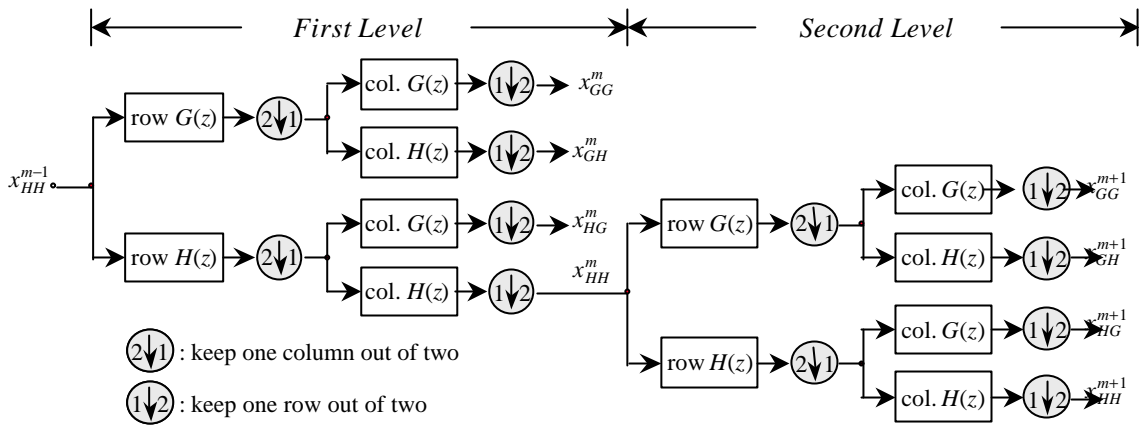


Fig 1 Two resolution levels of separate implementation for 2-D DWT decomposition.

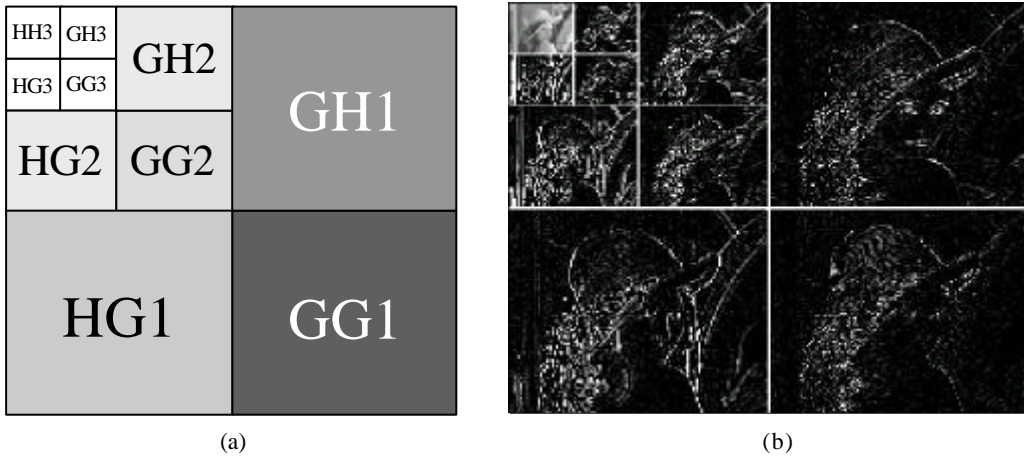


Fig 2 (a) Octave-band decomposition of an image. (b) Lena-image decomposition.

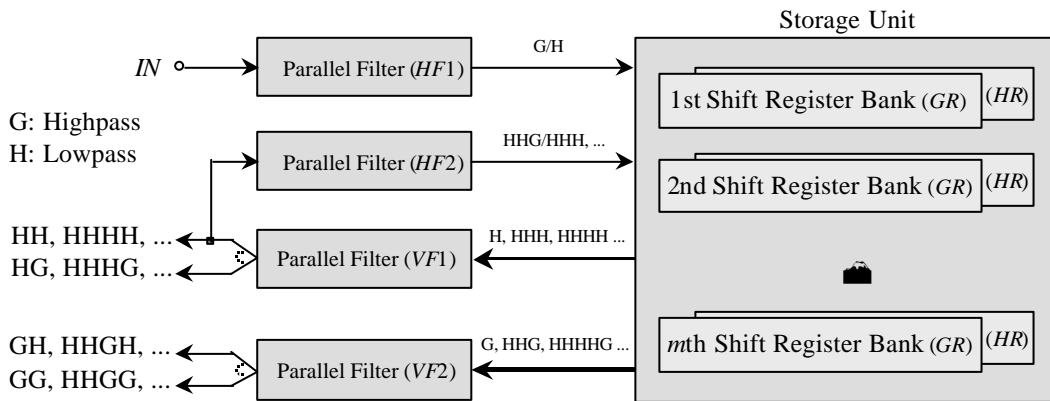


Fig 3 Proposed architecture for 2-D DWT.

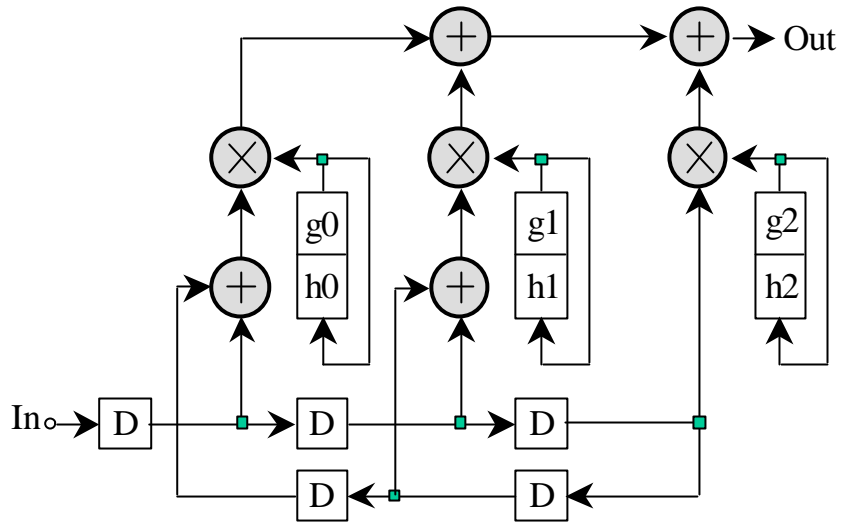


Fig 4 Five-tap high/low filter structure.

Table 1 A three-level computation schedule table.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	φ
0		H^1		H^1		H^1		H^1		H^1		H^1		H^1		H^1	φ
1		H^1	V^1	H^1	V^1	H^1	H^2 V^1	H^1	V^1	H^1	H^2 V^1	H^1	V^1	H^1	H^2 V^1	H^1	φ
2	V^1	H^1	H^2	H^1		H^1		H^1		H^1		H^1		H^1		H^1	φ
3		H^1	V^1	H^1	V^1	H^1	H^2 V^1	H^1 V^2	V^1	H^1	H^2 V^1	H^1 V^2	H^3 V^1	H^1	H^2 V^1	H^1 V^2	φ
4	V^1	H^1	H^2	H^1	H^3	H^1		H^1		H^1		H^1		H^1		H^1	φ
5		H^1	V^1	H^1	V^1	H^1	H^2 V^1	H^1	V^1	H^1	H^2 V^1	H^1	V^1	H^1	H^2 V^1	H^1	φ
6	V^1	H^1	H^2	H^1		H^1		H^1		H^1		H^1		H^1		H^1	φ
7		H^1	V^1	H^1	V^1	H^1	H^2 V^1	H^1 V^2	V^1	H^1	H^2 V^1	H^1 V^2	H^3 V^1	H^1 V^3	H^2 V^1	H^1 V^2	φ
\dots	\dots	\dots	\dots	\dots	\dots	\dots	\dots	\dots	\dots	\dots	\dots	\dots	\dots	\dots	\dots	\dots	φ
N	V^1		H^2	V^2	H^3	V^3											

H^i : The i th level of horizontal filterings.

V^i : The i th level of vertical filterings.

Table 2 Comparisons of various 2-D biorthogonal DWT architectures for even-tap filters.

Architectures	MAC's	Multipliers	Adders	Registers	Period
Direct Approach [7]	L	-	-	N^2	$4N^2$
Systolic-Parallel [7]	$2L$	$4L$	$4L$	$2NL + 4N$	$N^2 + N$
Parallel Filter [8]	-	$4L$	$4(L-1)$	$2NL + N$	$N^2 + N$
Systolic Array [13]	$3L$	-	-	$2NL$	$N^2 + N$
Semi-Systolic [14]	-	$3L$	$3(L-1)$	$29N$, for $L=6$	$N^2 + N$
Parallel Pipelined [15]	$5L/2$	-	-	$2N$	$N^2 + N$
Semi-Recursive [17]	-	$7L/2$	$4(L-1)$	$4NL$	$N^2 + N$
Ours	-	$15L/8$	$4(L-1)$	$2NL - 2N$	$\gg N^2$