






High-Speed Wavelet Image Processing Using the Winograd Method

N. N. Nagornov¹ , N. F. Semyonova¹ , and A. S. Abdulsalyamova² 

¹ North-Caucasus Federal University, Stavropol, Russia
sparta1392@mail.ru

² North-Caucasus Center for Mathematical Research, Stavropol, Russia

Abstract. Wavelets are actively used for solving of image processing problems in various fields of science and technology. Modern imaging systems have not kept pace with the rapid growth in the amount of digital visual information that needs to be processed, stored, and transmitted. Many approaches are being developed and used to speed up computations in the implementation of various image processing methods. This paper proposes the Winograd method (WM) to speed up the wavelet image processing methods on modern microelectronic devices. The scheme for wavelet image filtering using WM has been developed. WM application reduced the computational complexity of wavelet filtering asymptotically to 72.9% compared to the direct implementation. An evaluation based on the unit-gate model showed that WM reduces the device delay to 66.9%, 73.6%, and 68.8% for 4-, 6-, and 8-tap wavelets, respectively. Revealed that the larger the processed image fragments size, the less time is spent on wavelet filtering, but the larger the transformation matrices size, the more difficult their compilation and WM design on modern microelectronic devices. The obtained results can be used to improve the performance of wavelet image processing devices for image compression and denoising. WM hardware implementation on a field-programmable gate arrays and an application-specific integrated circuits to accelerate wavelet image processing is a promising direction for further research.

Keywords: Wavelet Transform · Digital Filtering · Group Pixel Processing · Computational Complexity · High-Speed Calculations · High-Performance Computing

1 Introduction

Wavelets are actively used for solving image processing problems in various fields of science and technology such as denoising [1], color image processing [2], video analysis [3]. However, modern imaging systems have not kept pace with the rapid growth in the amount of digital visual information that needs to be processed, stored, and transmitted. Many approaches are being developed and used to speed up computations in the implementation of various image processing methods. The authors of [4] focus on the evolution and application of various hardware architectures. The fast decomposition algorithms based on a different representation called product convolution extension has

been proposed in [5]. This decomposition can be efficiently estimated by assuming that multiple operator impulse responses are available. The new simple adjacent sum method is developed in [6] for multidimensional wavelet constructing. This method provides a systematic way to build multidimensional non-separable wavelet filter banks from two 1D low-pass filters, one of which is interpolating to increase image processing speed. The authors of [7] describe an asymmetric 2D Haar transform and extend it to wavelet packets containing an exponentially large number of bases. A basis selection algorithm is also proposing for optimal basis finding in wavelet packets. Various modern GPU optimization strategies for the discrete wavelet transform implementation such as the use of shared memory, registers, warp shuffling instructions, and parallelism at the level of threads and instructions are presented in [8]. A mixed memory structure for the Haar transform is proposed in which a multilevel transform can be performed with a single launch of the combined kernel. The paper [9] proposes a new algorithm for 2D discrete wavelet transform of high-resolution images on low-cost visual sensors and nodes of the Internet of things. The reduction in computational complexity and power consumption compared to modern low-memory 2D discrete wavelet transform methods are the main advantages of the proposed segmented modified fractional wavelet filter. However, all of these methods are based on pixel-by-pixel image processing. The Winograd method (WM) reduce image processing time due to group pixel processing. The processed image is assembled from fragments of a certain size which reduces the multiplications number by increasing the additions number.

The purpose of this paper is to accelerate wavelet image processing using WM on modern microelectronic devices.

2 Wavelet Image Processing Using the Direct Implementation and the Winograd Method

Wavelet filtering using direct implementation (DI) has the form

$$I_2(x) = \sum_{i=0}^{f-1} I_1(x-i)K(i), \quad (1)$$

where I_1 and I_2 are the original and processed 2D images, respectively, x is the row number of the pixel processed by f -tap wavelet filter K . The wavelet transform extracts local information about the signal in both frequency and time. High computational complexity is a significant disadvantage of this transform. The scheme of 1D wavelet filtering of an image fragment using DI is shown in Fig. 1a, where S_I is the original image fragment, L and H are the low- and high-pass wavelet filters, P_A and P_D are the processed image pixels with approximate and detailing image information, respectively.

Image filtering using WM in matrix form [10] can be presented as

$$Z = A^T \left((GK) \odot (B^T S) \right), \quad (2)$$

where: Z is the processed image fragment of size $z \times 1$; K is the wavelet filter of size $f \times 1$; S is the original image fragment of size $s \times 1$, where $s = z + f - 1$; A^T , G , B^T are the transformation matrices of sizes $z \times s$, $s \times f$, $s \times s$, respectively; \odot is

the element-wise matrix multiplication. Algorithms for matrices A^T , G , B^T obtaining are described in [11]. WM is denoted as $F(z, f)$. Digital filtering is performed on two computational channels corresponding to low- and high-frequency wavelet filters during wavelet image processing. The products of GL and GH are calculated in advance when using a specific wavelet. The product of S and the transformation matrix B^T can be computed before splitting the calculations into two channels because does not depend on the wavelet choice. Next, the element-wise multiplications $B^T S$ by GL and GH and the products of the obtained results with the transformation matrix A^T are performed over two computational channels. The scheme of 1D wavelet filtering of an image fragment using WM is shown in Fig. 1b, where S is the original image fragment, L and H are the low- and high-frequency wavelet filters, B^T , A^T , G are the transformation matrices, S_A and S_D are the processed image fragments with approximate and detailing image information, respectively.

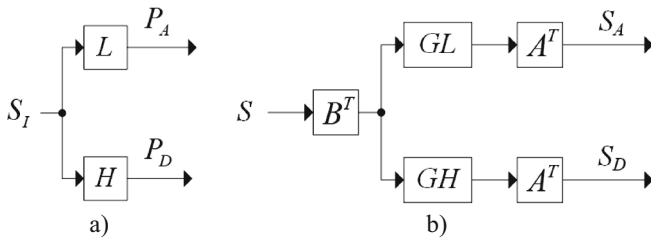


Fig. 1. The schemes of 1D wavelet filtering of an image fragment using: a) the direct implementation; b) the Winograd method

The results of increasing the speed of wavelet image processing using MW are presented below.

3 Acceleration of Wavelet Image Processing Using the Winograd Method

The computational complexity in time of wavelet filtering using WM $F(z, f)$ depends on the z and f and on the choice of points $s_0, s_1, \dots, s_{n-2}, s_{n-1}$. These values determine the form of transformation matrices A^T, G, B^T . The set of the Lagrange polynomial points $L = 0, 1, -1, 2, -2, 4, -4, \dots, 2^l, -2^l, 2^{l+1}, -2^{l+1}, \dots, \infty$ was used to construct the Vandermonde matrix V and matrices A^T, G, B^T [11]. The cases of using 4-, 6-, and 8-tap wavelets and processing of the original image fragments with size $z = 2, 3, 4, 5, 6, 7$ are considered. Table 1 is based on transformation matrices and contains the counting results of the multiplications and additions number required for wavelet filtering of images using DI and WM. The table values are obtained as follows.

1. The DI multiplications number is equal to the wavelet filters coefficients number.
2. The WM multiplications number is equal to twice the number of the processed image fragment pixels.

3. The DI additions number is equal to the number 2 less than the multiplications number.
4. The WM basic additions number is equal to the additions number of nonzero elements of matrices A^T (twice) and B^T by rows.
5. The WM complementary additions number is equal to the sum of the matrix element units in binary notation reduced by 1 for all elements of matrices A^T (twice) and B^T .
6. The total additions number is equal to the sum of basic and complementary additions.
7. WM receives several pixel values of the processed image in one iteration. Obtaining pixel brightness value requires the entire iteration as well as obtaining the entire fragment. Introduce the pixel specific value (PSV) for a correct comparison of the methods computational complexity. PSV is calculated as a quotient of the required operations number (multiplications or additions) divided by the number of pixels in the processed image fragment.

Table 1 shows that the greatest reduction in the specific weight of a pixel by multipliers is observed for 8-tap wavelet using WM $F(6, 8)$. The computational complexity decreases asymptotically by 72.9% compared to DI. The asymptotic estimate does not take into account addition operations since their complexity is an order of magnitude less than multiplication. This assessment is predominantly theoretical and may have a low correlation with the results obtained in the design of wavelet image processing devices in practice. Therefore, the unit-gate model (UGM) was used to calculate the operating time of a microelectronic device. UGM is a method for theoretical evaluation of device characteristics based on counting the number of the basic logical elements “and”, “or” [12]. The response time of one such element will be taken as a conventional unit (CU). Describe the principles of performing calculations in the theoretical estimation of the wavelet filtering devices delay according to the schemes in Fig. 1a and Fig. 1b for DI and WM, respectively. All multiplications are performed in parallel when using both methods.

Matrix multiplication operations can be replaced by shift and addition operations using the B^T and A^T matrices. The number of ones in the number binary representation for each element of the matrices A^T and B^T was calculated to determine the terms number in the rows of these matrices (Table 2). The products GL and GH are performed a priori. The products $B^T S$ on GL and GH are realized by element-wise multiplications. Multiplications and additions are implemented using a generalized multiplier (GM) and a multi-operand adder (MOA), respectively [13]. The delays of GM and MOA for k -bit numbers on computing devices are $6.8 \log_2 N + 2 \log_2 k + 4$ and $8.8 \log_2 k + 4$, respectively [14], where N is the largest number of elements in rows of matrices A^T and B^T , k is the image color depth and the coefficients bit-width of used wavelet filters. The calculations are performed for $k = 8$. The results of the device delay evaluation for wavelet image processing using DI and WM are presented in Table 2.

Table 1. The number of additions and multiplications in wavelet filtering of an image fragment using the direct implementation and the Winograd method

Tap	Method	Fragment pixels	For each fragment			Pixel specific value			
			Multiplications		Total	Additions		Total	
			Basic	Complementary		Basic	Complementary		
4	Direct	1	8	0	6	8.0	6.0	0.0	6.0
	$F(2,4)$	2	10	1	24	5.0	11.5	0.5	12.0
	$F(3,4)$	3	12	2	40	4.0	12.7	0.7	13.3
	$F(4,4)$	4	14	14	77	3.5	15.8	3.5	19.3
	$F(5,4)$	5	16	20	110	3.2	18.0	4.0	22.0
6	Direct	1	12	0	10	12.0	10.0	0.0	10.0
	$F(2,6)$	2	14	14	61	7.0	23.5	7.0	30.5
	$F(3,6)$	3	16	20	90	5.3	23.3	6.7	30.0
	$F(4,6)$	4	18	64	169	4.5	26.3	16.0	42.3
	$F(5,6)$	5	20	84	222	4.0	27.6	16.8	44.4
8	$F(6,6)$	6	22	197	382	3.7	30.8	32.8	63.7
	$F(7,6)$	7	24	248	478	3.4	32.9	35.4	68.3
	Direct	1	16	0	14	16.0	14.0	0.0	14.0
	$F(2,8)$	2	18	59	140	9.0	40.5	29.5	70.0
	$F(3,8)$	3	20	84	194	6.7	36.7	28.0	64.7
8	$F(4,8)$	4	22	197	350	5.5	38.3	49.3	87.5
	$F(5,8)$	5	24	248	442	4.8	38.8	49.6	88.4
	$F(6,8)$	6	26	505	756	4.3	41.8	84.2	126.0

Table 2. UGM-based evaluation results of the device delay for wavelet processing of 8-bit image using the direct implementation and the Winograd method

Tap	Method	Fragment pixels	The largest number of elements in a matrix row		Processing time according UGM	
			A^T	B^T	For each fragment	Pixel specific value
4	Direct	1	–	–	55.0	55.0
	$F(2, 4)$	2	4	4	78.6	39.3
	$F(3, 4)$	3	5	4	80.8	26.9
	$F(4, 4)$	4	6	9	90.5	22.6
	$F(5, 4)$	5	7	8	90.9	18.2
6	Direct	1	–	–	59.0	59.0
	$F(2, 6)$	2	6	9	90.5	45.3
	$F(3, 6)$	3	7	8	90.9	30.3
	$F(4, 6)$	4	8	18	100.2	25.0
	$F(5, 6)$	5	9	16	100.2	20.0
	$F(6, 6)$	6	10	34	108.6	18.1
	$F(7, 6)$	7	11	32	108.9	15.6
8	Direct	1	–	–	61.8	61.8
	$F(2, 8)$	2	8	18	100.2	50.1
	$F(3, 8)$	3	9	16	100.2	33.4
	$F(4, 8)$	4	10	34	108.6	27.2
	$F(5, 8)$	5	11	32	108.9	21.8
	$F(6, 8)$	6	12	58	115.6	19.3

The following conclusions are drawn based on the results in Table 2.

1. WM reduced the device delay of wavelet image processing to 66.9%, 73.6%, and 68.8% for 4-, 6-, and 8-tap wavelets, respectively, compared DI according to UGM.
2. The larger the processed image fragments size z , the less time is spent on wavelet filtering, but the larger the transformation matrices size, the more difficult their compilation and WM design on modern microelectronic devices.
3. The greatest reduction in device delay with an increase in the size of the resulting image fragments processed using WM is achieved at $z = 2$ and $z = 3$ according to UGM. For example, the device delay is reduced by $55.0 - 39.3 = 15.7$ CU and $39.3 - 26.9 = 12.4$ CU at $z = 2$ and $z = 3$, respectively, the device delay is reduced by 4.3 CU and 4.5 CU at $z = 4$ and $z = 5$, respectively, for 4-tap wavelet according to UGM.

4 Conclusion

The scheme for 1D wavelet image processing using WM has been developed. A comparative analysis of the image filtering time with DI was carried out. WM reduced the computational complexity of wavelet image processing asymptotically to 72.9% depending on the size of the filters used and fragments of the processed image. WM reduced the device delay of wavelet image processing to 66.9%, 73.6%, and 68.8% for 4-, 6-, and 8-tap wavelets, respectively, according to UGM. The larger the processed image fragments size z , the less time is spent on wavelet filtering, but the larger the transformation matrices size, the more difficult their compilation and WM design on modern microelectronic devices. The obtained results can be used to improve the performance of wavelet image processing devices for image compression and denoising. WM hardware implementation on FPGAs and ASICs to accelerate wavelet image processing is a promising direction for further research.

Acknowledgments. The work was supported by Russian Science Foundation, project № 22-71-00009.

References

1. Wu, Y., Gao, G., Cui, C.: Improved wavelet denoising by non-convex sparse regularization under double wavelet domains. *IEEE Access* **7**, 30659–30671 (2019)
2. Souillard, R., Carré, P.: Elliptical monogenic wavelets for the analysis and processing of color images. *IEEE Trans. Sig. Process.* **64**, 1535–1549 (2016)
3. Chen, Y., Li, D., Zhang, J.Q.: Complementary color wavelet: a novel tool for the color image/video analysis and processing. *IEEE Trans. Circuits Syst. Video Technol.* **29**, 12–27 (2019)
4. Alcaín, E., et al.: Hardware architectures for real-time medical imaging. *Electronics* **10**, 3118 (2021)
5. Escande, P., Weiss, P.: Fast wavelet decomposition of linear operators through product-convolution expansions. *IMA J. Numer. Anal.* **42**, 569–596 (2022)
6. Hur, Y., Zheng, F.: Prime coset sum: a systematic method for designing multi-d wavelet filter banks with fast algorithms. *IEEE Trans. Inf. Theory*, 7565569 (2016)
7. Ouyang, W., Zhao, T., Cham, W.K., Wei, L.: Fast full-search-equivalent pattern matching using asymmetric haar wavelet packets. *IEEE Trans. Circuits Syst. Video Technol.* **28**, 819–833 (2018)
8. Quan, T.M., Jeong, W.K.: A fast discrete wavelet transform using hybrid parallelism on GPUs. *IEEE Trans. Parallel Distrib. Syst.* **27**, 3088–3100 (2016)
9. Tausif, M., Khan, E., Hasan, M., Reisslein, M.: SMFrWF: segmented modified fractional wavelet filter: fast low-memory discrete wavelet transform (DWT). *IEEE Access* **7**, 84448–84467 (2019)
10. Winograd, S.: *Arithmetic Complexity of Computations* (1980)
11. Lyakhov, P., Abdulsalyamova, A., Semyonova, N., Nagornov, N.: On the computational complexity of 2D filtering by Winograd method. In: 2022 11th Mediterranean Conference on Embedded Computing (MECO), pp. 1–4 (2022)

12. Zimmermann, R.: Binary adder architectures for cell-based VLSI and their synthesis. Hartung-Gorre (1998)
13. Parhami, B.: Computer Arithmetic: Algorithms and Hardware Designs. Oxford University Press, Oxford (2010)
14. Lyakhov, P., Valueva, M., Valuev, G., Nagornov, N.: A method of increasing digital filter performance based on truncated multiply-accumulate units. *Appl. Sci.* **10** (2020)