Computational Restructuring: Rethinking Image Processing using Memristor Crossbar Arrays

Baogang Zhang, Necati Uysal and Rickard Ewetz
Department of Electrical and Computer Engineering, University of Central Florida, Orlando FL, USA
baogang.zhang@knights.ucf.edu, necati@knights.ucf.edu, rickard.ewetz@ucf.edu

***Abstract*—Image processing is a core operation performed on billions of sensor-devices in the Internet of Things (IoT). Emerging memristor crossbar arrays (MCAs) promise to perform matrix-vector multiplication (MVM) with extremely small energy-delay product, which is the dominating computation within the two-dimensional Discrete Cosine Transform (2D DCT). Earlier studies have directly mapped the digital implementation to MCA based hardware. The drawback is that the series computation is vulnerable to errors. Moreover, the implementation requires the use of large image block sizes, which is known to degrade the image quality. In this paper, we propose to restructure the 2D DCT into an equivalent single linear transformation (or MVM operation). The reconstruction eliminates the series computation and reduces the processed block sizes from NxN to $\sqrt{N}$x$\sqrt{N}$. Consequently, both the robustness to errors and the image quality is improved. Moreover, the latency, power, and area is reduced with 2X while eliminating the storage of intermediate data, and the power and area can be further reduced with up to 62% and 74% using frequency spectrum optimization.**

I. Introduction

Image and video compression is performed by transforming an image from the spatial domain into the frequency domain using the two-dimensional Discrete Cosine Transform (2D DCT) [7], [16]. Despite noteworthy efforts to accelerate image compression with algorithm innovations (as the Fast Fourier Transform [3], [15]) and custom digital hardware implementations [9], [14], the compression is still the bottleneck for real-time image and video processing systems [2], [8]. The 2D DCT for an image block involves performing a matrix-matrix-matrix multiplication. Due to promises of matrix-vector multiplication (MVM) with significant improvements in energy-delay product [4], [5], mixed analog-digital computing using memristor crossbar arrays (MCAs) has emerged as an appealing solution to accelerate 2D DCT [6], [10], [11].

The images obtained while performing image compression using digital and MCA based hardware are shown in Figure 1. In earlier studies [6], [10], [11], the two main explanations for the degraded image quality are: (i) Small errors in the output of the first matrix-matrix multiplication are amplified by the second matrix-matrix multiplication. (ii) Large block sizes are used to obtain the performance benefits associated with using MCAs with large dimensions [4]. However, small block sizes of 8x8 to 16x16 are essential to attaining high image quality [7]. Consequently, it is difficult (or impossible) to achieve high image quality when directly mapping the 2D DCT computation to MCA hardware.

Research was supported by NSF awards CCF-1755825 and CNS-1908471.

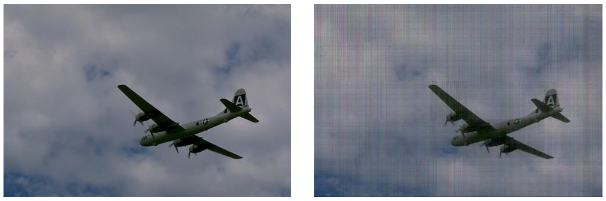

(a) (b)

Fig. 1. Image compression using (a) digital hardware and (b) MCAs. The MCAs have dimensions 64x128 and parameters as in [6], [10], [11].

In this paper, we propose to restructure the 2D DCT computation into an equivalent single linear transformation. The advantages of the reconstruction are, as follows:

- Compared with in [6], [10], [11], the number of MVM operations is reduced from $2N$ to N and no intermediate data is required to be stored.
- The robustness to variations is natively improved by eliminating the amplification of errors associated with performing two matrix-matrix multiplications in series.
- The reconstruction allows DCT matrices with dimensions NxN to process block sizes of $\sqrt{N}$x$\sqrt{N}$, and the performance benefits of using large MCAs are attained while utilizing traditional (small) block sizes.
- The reconstruction enables frequency spectrum optimization, which involves computing only a subset of the frequency coefficients. The optimization allows MCAs smaller dimensions to be used, which translates into power and area improvements.

The remainder of the paper is organized as follows: preliminaries are provided in Section II. The proposed 2D DCT reconstruction and frequency spectrum optimization is presented in Section III and Section IV. Experimental results are provided in Section V. The paper is concluded in Section VI.

II. Preliminaries

In this section, we review JPEG image compression using 2D DCT, how MCAs can accelerate MVM, and summarize the limitations of the previous work.

A. Image compression using 2D DCT

The fundamental steps of JPEG compression are illustrated in Figure 2 [13]. The first step is to partition the input image I into 8x8 image blocks X. Second, each image block X is converted into the frequency domain by applying the 2D DCT, i.e., $C = DXD'$, where C is a matrix of the frequency coefficients of X. D is the standard 2D DCT matrix. Third, the frequency coefficients are divided by each corresponding entry in a quantization table. The quantization is followed by zig-zag reordering, entropy encoding, and Huffman encoding. Next, a file is created that contains the compressed image and the encoding scheme. Uncompression is performed by reversing the process. The bottleneck of the overall flow is the 2D DCT, i.e., the computation of DXD' for each image block.

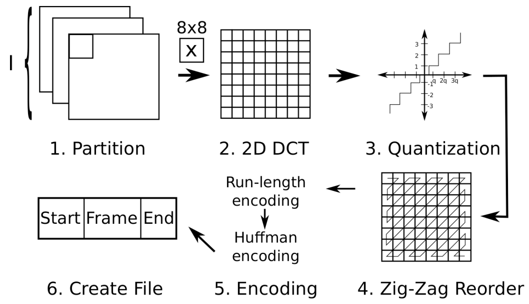


Fig. 2. Review of JPEG image compression [13].

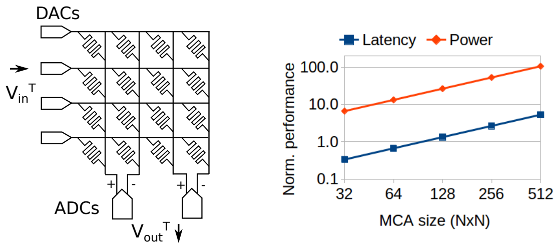
In this paper, the quality of the compression is measured using mean squared errors (MSE), as follows:

$$MSE(I,\hat{I}) = \frac{1}{PQ}\sum_{p=1}^{N}\sum_{q=1}^{M}(I_{pq}-\hat{I}_{pq}), \qquad (1)$$

where $\hat{I}$ is the original reference image with dimensions PxQ. I is the image obtained after $\hat{I}$ has been compressed and uncompressed using the flow in Figure 2.

B. Matrix-vector Multiplication using MCAs

An MCA consists of wordlines and bitlines with a memristor in each cross-point, which is shown in Figure 3(a). Analog matrix-vector is performed by passing an input vector v_{in} to the wordlines and recording an output vector v_{out} from the bitlines. The input and output voltages are converted between the digital/analog and analog/digital domain using DAC and ADC, respectively. As conductance values cannot be negative, the common differential pair approach is used to represent negative matrix values, i.e., a NxN matrix is represented using an Nx$2N$ MCA.



(a) (b)

Fig. 3. (a) MCA for MVM. (b) Normalized performance of MCA hardware vs digital hardware [4].

The advantage of leveraging MCAs is that the computation is orders of magnitude more efficient than using digital hardware, which is shown in Figure 3(b). The limitation is that the MVM is vulnerable to errors that are introduced by the array parasitics, analog variations, and the DACs and ADCs.

C. Previous Work and its Limitations [6], [10], [11]

In [6], [10], [11], 2D DCT was performed directly to the MCA based hardware, and $2N$ MVM operations are used to compute the frequency coefficients C of an image block X, which is illustrated in Figure 4.

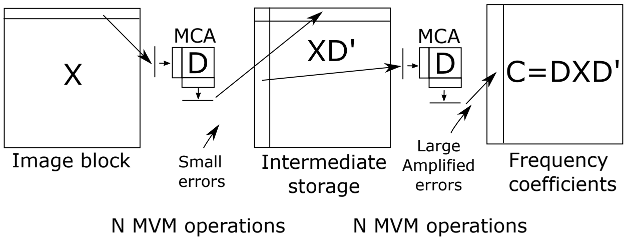


Fig. 4. Review of direct mapping in [6], [10], [11].

The two main limitations are: (i) The series matrix-matrix multiplication is inherently sensitive to variations. Small errors introduced in the first matrix-matrix multiplication are amplified into large errors by the second matrix-matrix multiplication. Due to the inherent presence of errors and variations within analog computing, it is impossible to achieve high image quality [6], [10], [11]. (ii) Large MCAs have to be utilized to gain performance advantages (power and latency) over digital implementations.

III. Proposed Reconstructed 2D DCT

A. Overview

In this paper, we propose to overcome the two aforementioned challenges based on reconstructing the 2D DCT computation into an equivalent linear transformation, which can be computed using a single MVM operation. It is easy to understand that such a transformation exists because two linear transformations are still a linear transformation. Consequently, X and C can be decomposed into vector form (x) and (c). Given the input vector x, the vector c is computed using a linear transformation as $c = \widetilde{D}x$, where $\widetilde{D}$ is a reconstructed 2D DCT matrix. The advantages of the reconstruction are: (i) There is no amplification of errors as the series computation is circumvented. (ii) Using MCAs with the exact same dimensions, the processed block size is reduced from NxN to $\sqrt{N}$x$\sqrt{N}$. (iii) The number of MVM operations required to process an image block of size NxN is reduced from $2N$ to N. (iv) The reconstruction opens-up new dimensions for optimization as each frequency coefficient in C is computed using a single row in $\widetilde{D}$.

B. Proposed Reconstruction

The mapping of 2D DCT to MCA hardware using the proposed reconstruction is shown in Figure 5. The figure shows that the image block X and the corresponding frequency representation C are divided into subblocks of size $\sqrt{N}$x$\sqrt{N}$, i.e., for a total of $N = \sqrt{N}$x$\sqrt{N}$ subblocks.

Let the image and frequency blocks respectively be denoted X_{ij} and C_{ij} with $1 \leq i \leq \sqrt{N}$ and $1 \leq j \leq \sqrt{N}$. Next, the subblocks are processed one-by-one into the corresponding

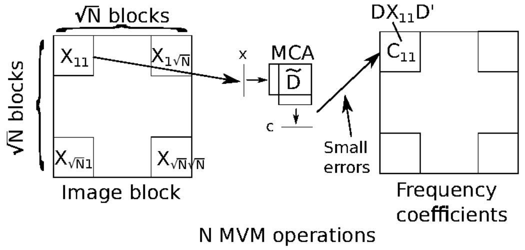


Fig. 5. Proposed 2D DCT computation using reconstructed DCT matrix.

frequency subblock, i.e., X_{ij} is processed into C_{ij}. Specifically, C_{ij} is obtained from X_{ij} by decomposing X_{ij} into a vector x row-wise (or column-wise). Next, the vector is passed to an MCA programmed with the matrix $\widetilde{D}$ to perform the computation $c = \widetilde{D}x$ efficiently in the analog domain, which is shown in the middle of Figure 5. The frequency block C_{ij} can be obtained from the output vector c by organizing the elements in c into a block format using the zig-zag pattern in Figure 2. In reality, there is no need to reorganize the vector c into the corresponding frequency subblock C_{ij}. The quantization table can instead be fused with $\widetilde{D}$ and run-length encoding can directly be applied to c. Consequently, the reconstruction can optionally eliminate step 3 and step 4 of the flow in Figure 2.

C. The reconstructed DCT matrix $\widetilde{D}$

The reconstructed 2D DCT matrix $\widetilde{D}$ is defined using a matrix $\overline{D}$ with the same dimensions. The matrices $\widetilde{D}$ and $\overline{D}$ are equivalent with respect to the ordering of the rows. $\overline{D}$ is defined with respect to a column decomposition of c into C whereas $\widetilde{D}$ is defined with respect to a zig-zag ordering. However, there is no simple closed-form expression for each element in $\widetilde{D}$. Therefore, we define an expression for $\overline{D}$ in Eq (2) and obtain $\widetilde{D}$ through reordering of the rows in $\overline{D}$. Let the element on row i and column j in $\overline{D}$ be denoted $\overline{D}_{ij}$, as follows:

$$\overline{D}_{ij} = a_p \cdot a_q \cdot cos[\frac{\pi p(2t+1)}{2N}] \cdot cos[\frac{\pi q(2r+1)}{2N}], \qquad (2)$$

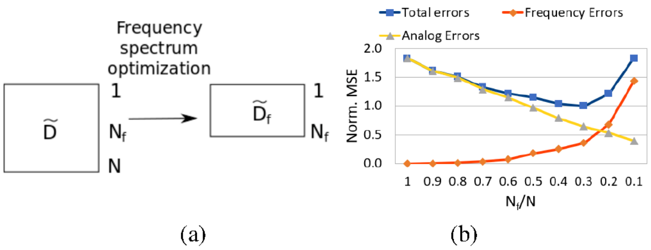
where $1 \leq i \leq N, 1 \leq j \leq N$. $q = i/N$ and $r = j/N$ where $/$ is integer division. $p = mod(i, N)$ and $t = mod(j, N)$ where mod is the modulus operator. The constant a_k is defined, as follows:

$$a_k = \begin{cases} \frac{1}{\sqrt{N}}, & k = 0, \\ \sqrt{\frac{2}{N}}, & k \neq 0. \end{cases}$$

Next, the rows in $\overline{D}$ are reordered to obtain $\widetilde{D}$. As mentioned earlier, the advantage of reordering the rows is that the subsequent zig-zag reordering step is automatically performed.

IV. Frequency Spectrum Optimization

In this section, we propose a frequency spectrum optimization technique. The technique involves computing only a subset N_f of the N and the frequency coefficients with respect to an image subblock X_{ij} with dimension $\sqrt{N}$x$\sqrt{N}$. Each row in the reconstructed DCT matrix $\widetilde{D}$ is used to compute a frequency coefficient in C. Consequently, the frequency spectrum optimization involves transforming $\widetilde{D}$ into a new matrix $\widetilde{D_f}$ with dimensions N_fxN, which is illustrated in Figure 6(a). Therefore, the dimension of the MCA used to perform the MVM operation can be correspondingly reduced, which translates into power and area savings.



(a) (b)

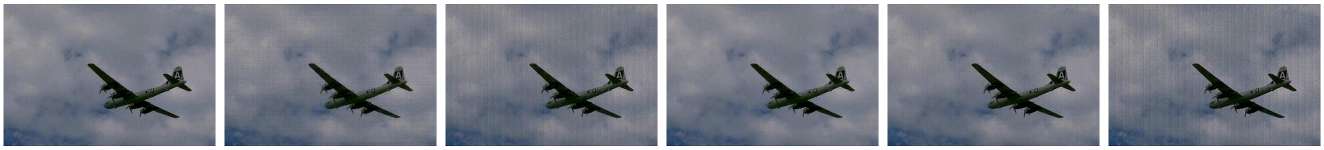
Fig. 6. (a) Reduction of $\widetilde{D}$ into $\widetilde{D}_f$ using by selecting N_f of the N frequency coefficients. (b) Trade-off between errors based on N_f/N for a reconstructed DCT matrix with dimensions 144x144. The figure shows that the smallest MSE is obtained when only a subset of the frequency spectrum is used.

In terms of image quality, *frequency errors* are introduced when only a subset of the frequency coefficients are computed. The frequency errors grow larger when the number of frequency coefficients (or basis functions) are reduced. Nevertheless, the image quality is gracefully degraded by ordering the frequency coefficients with respect to the zig-zag pattern and removing frequency terms starting from the tail-end. In contrast, the impact of analog errors is reduced when fewer frequency coefficients are computed. The explanation is that MCAs with smaller dimension introduce smaller analog errors because there is less IR-drop over the array parasitics [12].

The trade-off between frequency errors, analog errors, and total errors is shown as a function of N_f/N in Figure 6(b). The errors are measured in terms of MSE in Eq (1). The total errors are correlated with the image quality and are equal to the sum of the frequency errors and the analog errors. When the ratio N_f/N is reduced, the frequency errors are increased, which is illustrated with a red line in Figure 6(b). On the other hand, the analog errors are correlated with N_f/N, which is shown with a yellow line in Figure 6(b). Consequently, when N_f/N is increased, the total errors (blue line) are reduced until a turning point from were the errors start to increase rapidly. The turning point represents the optimal N_f that maximizes the image quality.

V. Experimental Results

The experimental results are obtained using a quad core 3.4 GHz Linux machine with 32GB of memory. The images in the evaluation are obtained from the Berkeley Segmentation Dataset and Benchmark Suite [1]. The images in the experimental results section are obtained by performing the compression using MCA hardware using the flow in Figure 2. The uncompression is performed by reversing the flow using digital hardware. The 2D DCT in MCA hardware is evaluated using circuit simulation with SPICE level accuracy while capturing array parasitics, programming accuracy, quantization



$\widetilde{D}$: 64x64 | 144x144 | 256x256 | $\widetilde{D}_f$: 64x39 | 144x44 | 256x52

MSE: 10.0 | 33.6 | 51.3 | 7.8 | 18.2 | 39.0

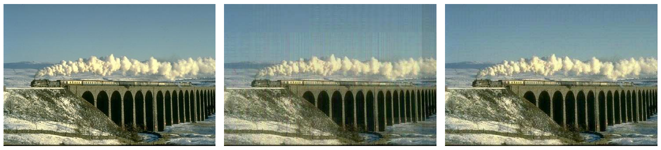
(a) Full frequency spectrum (b) Frequency spectrum optimization

Fig. 7. The images on the left (right) are obtained without (with) frequency spectrum optimization. The dimension of the reconstructed 2D DCT matrix and the MSE are shown below each figure.

errors in the domain interfaces, random telegraph noise (RTN), etc. The experimental setup has been proven to exhibit extremely high correlation with results obtained using hardware prototypes [5], [10]. We evaluate the 2D DCT reconstruction in Section V-A. The frequency spectrum optimization is evaluated in Section V-B.

A. Evaluation of Reconstruction

In Figure 8, we evaluate impact of the reconstruction of the image quality using MCAs with dimensions 64x128. The reference image is shown in Figure 8(a). The images obtained using the direct mapping in [6], [10], [11] and the proposed reconstruction are shown in (b) and (c) of Figure 8, respectively. The reference image is of high quality. In Figure 8(b), it can be observed that the image quality is degraded by the image compression in [6], [10], [11]. The image obtained after the proposed reconstruction in Figure 8(c) shows that the image quality is just slightly degraded compared with the reference image although MCA hardware is used. Moreover, the reconstruction improves the robustness to quantization errors introduced by the domain interfaces and random telegraph noise. Furthermore, the reconstruction improves power, area, and latency by 2X, as the number of MVM operations is reduced with 2X.



(a) (b) (c)

Fig. 8. (a) Reference image. (b) Image obtained with the image compression in [6], [10], [11]. (c) Image obtained using the proposed reconstruction.

B. Evaluation of Frequency Optimization

The frequency spectrum optimization is evaluated in terms of image quality in Figure 7. The figure shows that the image quality is gracefully degraded when MCAs with larger dimensions are utilized due to the IR-drop over the array parasitics. For reconstructed 2D DCT matrices $\widetilde{D}$ with dimensions 64x64, 144x144, and 256x256, the optimal $\hat{N}_f^*/N$ is determined to be 0.6, 0.3, and 0.2. The optimal ratios were obtained by evaluating a collection of 40 images using MCAs with different dimensions and selecting the ratio that minimized MSE in Eq (1). It is not surprising that the N_f^*/N ratio becomes smaller for MCAs with larger dimensions, as larger MCAs are more severely impacted by IR-drop over the array parasitics [12]. The figure shows that the frequency spectrum optimization reduces the MSE for the images in each column. Moreover, the frequency optimization reduces the power and area overhead with up to 61.5% and 74.3%, respectively.

In summary, it is highly advantageous to reconstruct the 2D DCT matrix and apply frequency spectrum optimization. Compared with in [6], [10], [11], the image quality is improved, the power is improved with 65%, 90%, and 95%, and the area is improved with 69%, 92%, and 97% for MCAs with 64, 144, and 256 inputs, respectively.

VI. Summary and Future Work

In this paper, a reconstruction of the 2D DCT matrix is proposed along with a frequency spectrum optimization technique. The techniques demonstrate (i) significant improvements in image quality, (ii) higher robustness to errors, (iii) notably smaller power, area, and latency compared with in previous studies. In our future work, we will investigate techniques for further improvement in the analog domain.

References

[1] https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/.

[2] O. Fialka and M. Cadik. Fft and convolution performance in image filtering on gpu. pages 609 – 614, 08 2006.

[3] Fusheng Zhang, Zhongxing Geng, and Wei Yuan. The algorithm of interpolating windowed fft for harmonic analysis of electric power system. *PWRD*, 16(2):160–164, April 2001.

[4] M. Hu et al. Dot-product engine for neuromorphic computing: Programming 1t1m crossbar to accelerate matrix-vector multiplication. DAC, pages 1–6, 2016.

[5] M. Hu et al. Memristor-based analog computation and neural network classification with a DPE. *Adv. Materials*, 30, 2018.

[6] M. Hu and J. P. Strachan. Accelerating discrete fourier transforms with dot-product engine. ICRC, pages 1–5, 2016.

[7] A. K. Jain. Image data compression: A review. *Proceedings of the IEEE*, 69(3):349–389, 1981.

[8] P. Karas and D. Svoboda. Algorithms for efficient computation of convolution. 2013.

[9] E. Konguvel and M. Kannan. A survey on fft/ifft processors for next generation telecommunication systems. *Journal of Circuits, Systems and Computers*, 27(03):1830001, 2018.

[10] C. Li et al. Analogue signal and image processing with large memristor crossbars. *Nature Electronics*, 1(1):52, 2018.

[11] C. Li et al. Large memristor crossbars for analog computing. ISCAS, pages 1–4, 2018.

[12] B. Liu et al. Reduction and ir-drop compensations techniques for reliable neuromorphic computing systems. ICCAD, pages 63–70, 2014.

[13] Z. Liu et al. Deepn-jpeg: A deep neural network favorable jpeg-based image compression framework. DAC, page 18, 2018.

[14] A. Pedram, J. McCalpin, and A. Gerstlauer. Transforming a linear algebra core to an fft accelerator. ASAP, pages 175–184, June 2013.

[15] C. Van Loan. *Computational Frameworks for the Fast Fourier Transform*. Society for Industrial and Applied Mathematics, 1992.

[16] G. K. Wallace. The jpeg still picture compression standard. *IEEE transactions on consumer electronics*, 38(1):xviii–xxxiv, 1992.