

Received January 30, 2022, accepted February 10, 2022, date of publication February 14, 2022, date of current version February 23, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3151361

# RNS-Based FPGA Accelerators for High-Quality 3D Medical Image Wavelet Processing Using Scaled Filter Coefficients

NIKOLAY N. NAGORNOV<sup>1</sup>, PAVEL A. LYAKHOV<sup>1,2</sup>,  
MARIA V. VALUEVA<sup>2</sup>, (Graduate Student Member, IEEE),  
AND MAXIM V. BERGERMAN<sup>2</sup>, (Graduate Student Member, IEEE)

<sup>1</sup>Department of Mathematical Modeling, North-Caucasus Federal University, 355017 Stavropol, Russia

<sup>2</sup>North-Caucasus Center for Mathematical Research, North-Caucasus Federal University, 355017 Stavropol, Russia

Corresponding author: Nikolay N. Nagornov (sparta1392@mail.ru)

This work was supported in part by the North-Caucasus Center for Mathematical Research through the Ministry of Science and Higher Education of the Russian Federation under Grant 075-02-2022-892 (research in Sections III-B, IV-B, and V-B), and in part by the Russian Science Foundation under Grant 21-71-00017 (research in the remaining sections).

**ABSTRACT** Medical imaging using different modalities has many problems. The main ones are low informativeness, various distortion noises, and a large amount of information. Fusion, denoising, and visual data compression are used to solve them in practice. Discrete wavelet transform is one way to implement various fusion, denoising, and compression methods for 2D and 3D medical image processing. Medical imaging systems produce increasingly accurate images with scanning technology and digital devices development. These images have improved quality using both higher spatial resolutions and color bit-depth. Processing a large volume of medical imaging data requires considerable resources and processing time. Modern wavelet-based devices for medical image processing do not meet the current performance demand. Hardware accelerators are being designed to solve this problem. This paper proposes new (field-programmable gate array) FPGA accelerators using wavelet processing (WP) with scaled filter coefficients (SFC) and parallel computing in residue number system (RNS) to improve the performance of high-quality 3D medical image WP systems. The computational complexity is reduced using the developed WP method with SFC and the proposed wavelet filter coefficients scaling algorithm. Parallel computing is organized in RNS using moduli sets of a particular type. Hardware implementation of 3D medical image WP using the proposed FPGA accelerators increases device performance by 2.89-3.59 times, increasing the hardware resources by 1.18-3.29 times compared to state-of-the-art solutions. The device performance improvement is achieved while maintaining high-quality 3D medical image processing in peak signal-to-noise ratio terms.

**INDEX TERMS** Medical image processing, discrete wavelet transform, scaled filter coefficients, residue number system, high-performance computing, hardware accelerator, field-programmable gate arrays.

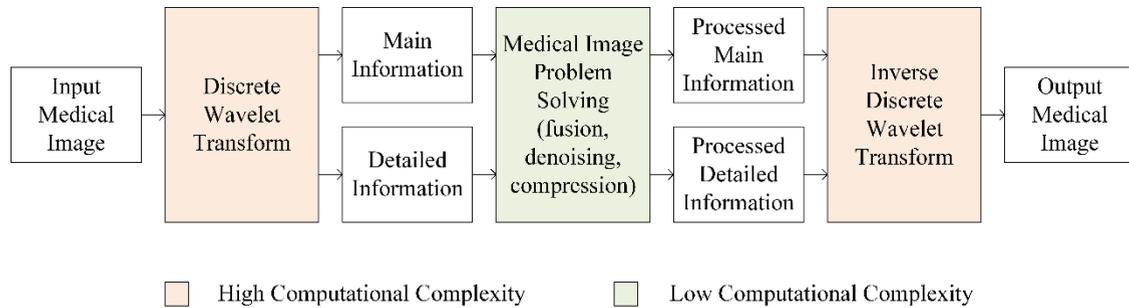
## I. INTRODUCTION

Various medical imaging modalities are used for diagnosis purposes and treatments. Magnetic resonance imaging (MRI) [1], X-ray [2], ultrasound [3], radionuclide [4], and optical [5] are the most common of them. Modern medical imaging systems have many problems associated with visual data processing. Low informativeness of 2D and 3D images often does not contain enough information for high-quality diagnosis is one of them. Multiple views combining

The associate editor coordinating the review of this manuscript and approving it for publication was Larbi Boubchir.

the same organ solve this problem in practice [6], [7]. The resulting image is more informative and facilitates perception by both humans and machines, increasing diagnostics' accuracy. Modalities MRI, computed tomography (CT), positron emission tomography (PET), single-photon emission computed tomography (SPECT), and their hybrids CT/PET, CT/SPECT, CT/MRI, MRI/PET, MRI/SPECT are used for brain diseases diagnosis [8], for example. Such combinations provide the complete description of the organ anatomical structure and follow up the organ cell behavior.

Medical image denoising is also an urgent problem for many imaging modalities [9]–[11]. Images are subject to



**FIGURE 1.** Medical image wavelet processing scheme.

noise that distorts information and interferes with medical diagnosis and disease treatment. Ultrasound imaging presents low-resolution images usually degraded by speckle noise requiring imaging techniques to improve image quality and proper medical diagnosis [12]. Noise affects the anatomical structure's analysis using CT or related morphological imaging modalities and impedes diagnosis [13]. MR images are inherently noisy, and thus, filtering methods are required to image quality improvement [14]. The PET reconstruction process includes inherent multiplicative noise, preventing visual data analysis [15]. Optical coherence tomography (OCT) noise in retinal imaging limits the human eye's structural features, such as retinal layer properties [16].

Scanning technologies are evolving. Digital devices are constantly improving. Medical imaging systems produce increasingly accurate images with improved quality using higher spatial resolutions and color bit-depth. Such improvements increase the amount of information that needs to be stored, processed, and transmitted. This process is significant for 3D scanning technology [17]. Medical imaging sets of 10-30 TB per patient are not unusual. Tomograph that registers images of 1196 by 500 pixels, 24 bits per pixel, and 20 frames per second creates a tomogram of about 20 GB in 10 minutes [18]. The retina OCT results may require more than 40 GB [19]. Virtual microscopy color image sets can often have a size exceeding 50 GB [20]. The medical system can generate data at speeds of up to 4 GB per second for several days with selective plane illumination microscopy [21]. Enormous computing resources are needed to process such massive amounts of data. A significant part of space is required to store them. High throughput is needed to transfer them. Thus, effective 3D image compression is an urgent problem of modern medical imaging systems [22]–[24].

In practice, various transforms are used for medical image fusion, denoising, and compression. Discrete Fourier transform (DFT) [25], [26], Hadamard transform (HT) [27], and discrete wavelet transform (DWT) [28], [29] are the most common of them. Both DFT and HT are widely used in the frequency domain, but the time domain characteristics disappear. The time and intensity levels after signal DFT or HT cannot be determined. The local properties of the signal become impossible. DWT solves these problems since

it obtains both signal frequency and time information [30]. Furthermore, DWT is the only one of the listed transforms that can solve any of the three outlined problems: low information content, noisiness, excessive storage space. DWT is a signal transform using a filter bank that convolutions the input data with wavelet filters. DWT translates the signal from a time representation into a time-frequency domain. Image DWT is performed by convolution with a pair of lowpass and highpass wavelet filters of the filter bank. These filters highlight main and detailed information, respectively. Images' wavelet processing (WP) is carried out in three main stages (Fig. 1). DWT extracts both main and detailed information from the input medical image. Manipulations over the main and detail coefficients allow solving denoising, compression, and fusion problems. Inverse DWT forms the output medical image from the processed main and detailed information. Image fusion is performed by basically combining main image information from different modalities [26]. Both denoising and image compression are primarily performed by detailed information manipulating [31], [32]. The first and third stages of medical image WP have high computational complexity. They are analyzed in this paper. WP means the sequence of DWT and inverse DWT in the rest of the paper. Processing a large volume of medical imaging data requires many resources and processing time. Modern wavelet-based devices of medical image processing do not meet the current performance demand [20], [21], [33]. The new medical visual data processing methods are being developed and implemented on modern high-performance hardware accelerators. The field-programmable gate array (FPGA) [34] and application-specific integrated circuit (ASIC) [35] are some of the main. These accelerators improve the speed and quality of the solution to the problems designated.

Data representation accuracy significantly impacts computational complexity. It is an important problem in modern hardware accelerator design [37]. Computational complexity is the main factor affecting device performance, hardware costs, and power consumption. The data representation accuracy does not receive significant attention when medical image WP is implemented in modern software packages. The wavelet filter coefficients have a bit-width of 16 or 32 due to peculiar modern general-purpose processor architecture.

The question of the data representation effective accuracy in wavelet image processing on modern hardware accelerators with a more flexible architecture is relevant. The accuracy selected directly affects operations complexity and computational speed [38]. Thus, the choice of the bitness of data representation in the memory of hardware devices is important. It should reduce computational complexity. The quality of image processing should remain high at the same time.

Parallel computing is the primary approach to the performance improvement of software and hardware devices today. However, positional number systems are traditionally used for computations organizing in microprocessor systems development. They do not allow its effective implementation because there are many inter-digit carries when arithmetic operations perform. The residue number system (RNS) is one of the most famous alternative numeric systems capable of parallel arithmetic operations [39]. RNS is a direct sum of finite rings. It operates with small residues instead of large-digit numbers processing in positional number systems, which opens up opportunities for high-speed parallel data processing at the arithmetic-logic level on modern microelectronic devices [40]. Furthermore, RNS uses arbitrary bitness calculations and opens up the possibility of image processing device detailed customization.

The goal of this paper is to increase the hardware device performance for 3D medical image WP. We scale wavelet filters coefficients to reduce the digital filtering computational complexity and use RNS for parallelizing the computations. The main contributions of this paper are as follows:

1. We develop the WP method with scaled filter coefficients (SFC), which reduces computational complexity, simplifies rounding operations, and uses integer calculations in fixed-point format for high-performance and area-efficient modern hardware accelerators implementation.
2. We propose the wavelet filter coefficients scaling algorithm that reduce the digital filtering computational complexity while maintaining high-quality medical image WP in terms of peak signal-to-noise ratio (PSNR).
3. We propose the 3D medical image WP implementation using RNS-based parallel computing with moduli sets of a particular type. Use special techniques for arithmetic operations performing in adders and significantly reduce latency across all hardware device calculation channels.
4. We propose the implementation of 3D medical image WP on FPGAs and a comparison with state-of-the-art solutions. We prove that the SFC reduces the hardware resources and processing time while maintaining high processing quality in PSNR terms, and parallel computing in RNS can increase hardware device performance.

The rest of the paper is organized as follows. Section II contains related work on 4 points: medical image processing using wavelets; hardware accelerators for medical image processing; wavelets accuracy in hardware devices; hardware digital filtering in the residue number system. Section III provides the proposed approach description to high-performance hardware WP of 3D medical images on 2 points: 3D images

WP method with SFC and wavelet filter coefficients scaling algorithm; RNS parallel computing and hardware architecture features. Section IV describes the software and hardware implementation of the proposed WP method and algorithm with SFC for 3D medical image processing using RNS parallel computing. Section V includes results, discussion, and future work. Section 6 presents the conclusion. The Application contains a numerical method for evaluating the calculation error of WP using SFC.

## II. RELATED WORK

This section is divided into 4 subsections. In subsection IIA, an overview of the wavelets used for image processing in medicine is provided. In subsection IIB, specialized hardware accelerator developments for medical image processing are described. Subsection IIC is devoted to works in which the wavelet accuracy representation in the hardware device memory is raised. Subsection IID presents a brief analysis of approaches to digital filtering hardware implementation methods in RNS.

### A. MEDICAL IMAGE PROCESSING USING WAVELETS

Medical imaging mainly uses wavelets to solve three important problems: fusion, denoising, and compression of medical images. The authors [26] proposed an efficient colorful Fourier ptychographic microscopy reconstruction method using multi-resolution wavelet-based fusion. Wavelet fusion method and algorithm for the fusion of intravascular ultrasound and OCT pullbacks to improve the use of those two types of medical images are proposed in [5]. Banik *et al.* [7] proposed a fusion-based polyp segmentation network using dual-tree wavelet transform. A feature extraction-based method using the improved Haar wavelet for grading cataract severity on retinal images is proposed in [28]. Rodrigues *et al.* [12] offered new wavelet and bilateral filtering methods for speckle noise reduction in medical ultrasound images. DWT-based deconvolution method for quantitative oncologic PET imaging was implemented by Rezaei *et al.* [4]. The authors [29] presented an iterative image reconstruction algorithm using undecimated wavelet transform for MRI denoising. An objective method for computing the optimum threshold value in the dual-tree complex WP based on real-time MR images denoising is proposed in [31]. Krishnaswamy [41] proposed an efficient image codec based on frequency domain transformation using lifting wavelet transform for efficient MRI compression. An algorithm for perceptually lossless volumetric CT image compression using fractional wavelet filtering is presented in [23]. A new enhanced embedded zerotree wavelet algorithm for medical image compression applications is presented in [32]. The authors [24] proposed a new compression strategy of MRI brain datasets using a DWT-based coder.

We can conclude that wavelets are widely used in medicine to process images of various modalities. However, medical imaging devices produce increasingly accurate images with improved quality, higher resolution, and increased color

depth. The need to modify existing and develop new methods and algorithms to improve existing and create new high-performance medical image processing systems is constantly growing. Specialized hardware accelerator designing discussed further is one approach to solving this problem.

### B. HARDWARE ACCELERATORS FOR MEDICAL IMAGE PROCESSING

The world scientific community develops and improves medical imaging methods to increase the speed of various problems solving. These methods are implemented on modern high-performance hardware microelectronic devices such as FPGA and ASIC. FPGAs are most common because they have a low manufacturing cost and can be reprogrammed. Ravi *et al.* [34] used FPGA accelerators to increase the maximum-likelihood expectation-maximization algorithm's speed for tomographic image reconstruction. The prototype of the first portable 3D digital ultrasonic back-end system implemented in FPGA is designed in [42]. The authors [43] presented the Gabor filter design on the FPGA accelerator for power-efficient and high-performance medical imaging applications. FPGA-based emulation of minimum variance distortionless response beamformer architecture for ultrasound imaging is proposed in [44]. A novel method for the urinary bladder non-invasively volume measuring based on the electrical impedance tomography principles and implemented on FPGA is presented in [25]. The authors [45] designed a novel OCT imaging portable processor implemented on energy-efficient and calibration-aware FPGA. ASICs consume less power and are faster than FPGAs but cannot be reprogrammed and are expensive to manufacture. A reconfigurable ultrasound transceiver ASIC for 3D carotid artery imaging is presented in [46]. Nadig *et al.* [47] described the investigation of the ASIC power consumption in PET systems. The authors [48] designed the Anger camera based on the silicon photomultiplier and custom ASIC for MRI scanners. Trigilio *et al.* [49] developed the ASIC for silicon photomultiplier readout targeting SPECT applications based on Anger-type detectors. A 2D X-ray detector using ASIC for fine pitch and high-energy resolution imaging spectroscopy is developed in [50].

We can conclude that the development and implementation of various methods for the performance of medical visual data processing devices are increasing based on hardware accelerators FPGA and ASIC. This paper proposes two new ideas: the data representation accuracy decreasing in the hardware devices memory, which does not lead to an image WP quality deterioration; a parallel computing organization in RNS. An analysis of related work on these ideas is presented in the following 2 subsections.

### C. WAVELETS ACCURACY IN HARDWARE DEVICES

The accuracy of the digital filters coefficients representation significantly impacts image processing quality and the digital filtering computational complexity. Computational complexity affects the time, hardware, and energy resources in the

hardware implementation of signal and image WP methods. Thus, the representation accuracy of the digital filter coefficients in the modern microelectronic devices memory is one of the important problems in the computer technology devices design. Few works address this problem. Experimental results from [51] suggested that 16-bit integers provide an acceptable computational range for representing the wavelet coefficients for integer-to-integer transforms at the 8-bit images processing. Larrotta *et al.* [52] proposed an FPGA architecture designed with the requirements of the biomedical signals compression model based on wavelet transform and run-length encoding using 16-bit wavelet filter coefficients. The authors [53] presented a new high-performance and low-area hardware configurable architecture for DWT using 16-bit wavelet filter coefficients. A new high-performance FPGA-based parallel architecture of inverse DWT with 16-bit wavelet coefficients is presented in [54]. The authors [38] investigated the design and implementation aspects of a multiplier-free 2-level WP using Daubechies wavelet with 12-bit coefficients and parallelized computations in the 3-modulo RNS for high-speed signal processing applications. The PSNR dependence of 2D images direct multilevel wavelet transform on the wavelet used, coefficients bit-width, and the processing levels number is revealed in [55].

The wavelet representation accuracy is determined empirically without any theoretical justification and only for specific cases in the analyzed works. The theoretical analysis of the calculation error dependence on the hardware device memory's wavelet representation accuracy is absent. The authors use various tools to organize computations when image WP methods are implemented. We propose using RNS-based calculations since it allows arbitrary bitness calculations and opens up the possibility of detailed image processing device customization. A brief analysis of approaches to the digital filtering methods hardware implementation in RNS is presented further.

### D. HARDWARE DIGITAL FILTERING IN THE RESIDUE NUMBER SYSTEM

RNS is a non-positional number system that represents numbers as a tuple of remainders after division by pairwise coprime moduli  $\{m_1, m_2, \dots, m_j\}$ , where  $GCD(m_{j_1}, m_{j_2}) = 1$  for  $j_1 \neq j_2$  and  $j$  is the number of RNS moduli. Any number  $A$  satisfying the condition  $0 \leq A < D$  can be uniquely represented in RNS as a set  $A = (a_1, a_2, \dots, a_j)_{RNS}$ , where  $a_j = |A|_{m_j}$  and  $D = \prod_{i=1}^j m_i$  is the dynamic range [56]. RNS dynamic range is divided into two roughly equal parts representing negative numbers. Therefore, any integer can be represented in RNS if it satisfies one of the two relations [57]:

$$\begin{aligned} -(D-1)/2 \leq A \leq (D-1)/2 & \quad \text{for odd } D, \\ -D/2 \leq A \leq D/2 - 1 & \quad \text{for even } D. \end{aligned}$$

Modular operations are defined as follows:

$$A * B = \left( |a_1 * b_1|_{m_1}, |a_2 * b_2|_{m_2}, \dots, |a_j * b_j|_{m_j} \right)_{RNS},$$

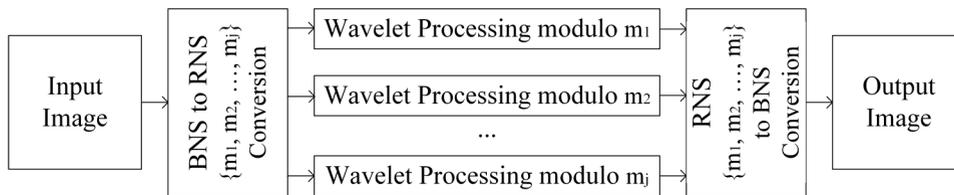


FIGURE 2. Wavelet processing of visual data in the residue number system.

where “\*” is the addition, subtraction, or multiplication operation. Number  $A = (a_1, a_2, \dots, a_j)_{RNS}$  conversion from RNS to BNS is based on the Chinese remainder theorem [57]:

$$A = \left\| \sum_{i=1}^j \left( \left\| D_i^{-1} \right\|_{m_i} a_i D_i \right) \right\|_D,$$

where:  $D_i = D/m_i$ ;  $\left\| D_i^{-1} \right\|_{m_i}$  is the  $D_i$ . multiplicative inverse.

Example. Let us have RNS with a set of  $j$  pairwise coprime moduli  $\{2, 3, 5, 7\}$ .  $GCD(2, 3) = GCD(2, 5) = GCD(2, 7) = GCD(3, 5) = GCD(3, 7) = GCD(5, 7) = 1$ . Then  $D = 2 \cdot 3 \cdot 5 \cdot 7 = 210$ . Let us have numbers  $A = 53$  and  $B = 131$ . Then  $A = (1, 2, 3, 4)_{RNS}$  and  $B = (1, 2, 1, 5)_{RNS}$ . Their sum  $A + B = (|1 + 1|_2, |2 + 2|_3, |3 + 1|_5, |4 + 5|_7)_{RNS} = (0, 0, 4, 2)_{RNS}$ . At the same time,  $A + B = 53 + 131 = 184 = (0, 0, 4, 2)_{RNS}$  if we use RNS for positive numbers.

Thus, BNS operations with a large capacity are reduced to RNS operations on numbers with a small capacity. The computations organization in RNS opens up the possibility for high-speed parallel data processing at the arithmetic-logic level. RNS is used for the hardware accelerators development in applications that mainly use addition, subtraction, and multiplication operations. Digital filtering is one such application widely used for various digital signal and image processing problem-solving. Safari et al. [58] implemented the power-performance enhancement of a 2D WP image processor using RNS and the static voltage scaling scheme. High-speed smoothing filter RNS-based architecture is proposed in [59]. A low-power RNS-enhanced arithmetic unit designing for constant-coefficient filtering is offered in [60]. The authors [61] investigated the design and implementation of finite impulse response filters on the RNS-based FPGA. A high-performance FPGA accelerator for the multiplier-free two-level WP implementation using RNS is proposed in [38]. Image filtering implementation method in RNS with replacing the computationally complex division operation by the multiplication and scaling is developed in [62]. A high-performance hardware implementation of digital filtering using truncated multiply-accumulate units in RNS is presented in [63]. FPGA simulation results show that the proposed RNS-based approach increases the digital filtering speed by about 4 times and reduces the hardware costs by 3 times compared to the binary number system (BNS). Hardware devices using RNS-based wavelet transform, which increase the performance of 3D image processing devices

by up to 2.70 times, are described in [64]. FPGA accelerator for signal processing and machine learning algorithms implementation with multiply-accumulate units in RNS is developed in [65]. A convolutional neural network architecture with RNS-based digital filters hardware implementation is designed in [66]. The convolutional layer implementation has shown that RNS calculations can reduce hardware costs by up to 37.78% and the image recognition time by up to 41.17% compared to the two’s complement implementation. RNS-based solution for the hardware cost reducing of digital filtering in a neural network convolutional layer is developed in [67]. RNS-based calculations in the neural network convolutional layer reduced device area by 32.6% compared to BNS-based approach by the hardware simulation results on FPGA. The approach scheme used in the analyzed works for RNS-based visual data WP is shown in Fig. 2. The processing is performed in three stages: the input data is converted from a BNS to an RNS with moduli set  $\{m_1, m_2, \dots, m_j\}$ ; image WP in RNS is performed; the processing results are converted from RNS with moduli set  $\{m_1, m_2, \dots, m_j\}$  to BNS and transmitted to the output. The RNS-based data transforming and processing methods and algorithms are detailed in [40].

We can conclude that parallel computing in RNS can improve the performance of the digital image processing hardware device. The proposed approach to the hardware FPGA accelerators development and implementation for 3D medical image WP is presented further.

### III. PROPOSED APPROACH

We propose a new approach to high-performance hardware WP of 3D medical images based on SFC and RNS parallel computing. A scheme of the proposed approach is shown in Fig. 3. Apriori information about both the modality (or modalities) of the 3D medical image (or images) formation and image processing problems arising is used in the proposed approach scheme (Fig. 3). This information includes the input image, image color depth, and the image processing quality criterion. The wavelet is selected based on the medical image processing problem (fusion, denoising, or compression). Preliminary calculations are carried out by wavelet filter coefficients scaling using the proposed algorithm. The algorithm is detailed in subsection IIIA. The obtained results were converted into RNS with moduli set  $\{m_1, m_2, \dots, m_j\}$ . WP of medical data in RNS is performed on FPGA using the well-known approach described in subsection IID. The high-performance WP method using SFC is presented

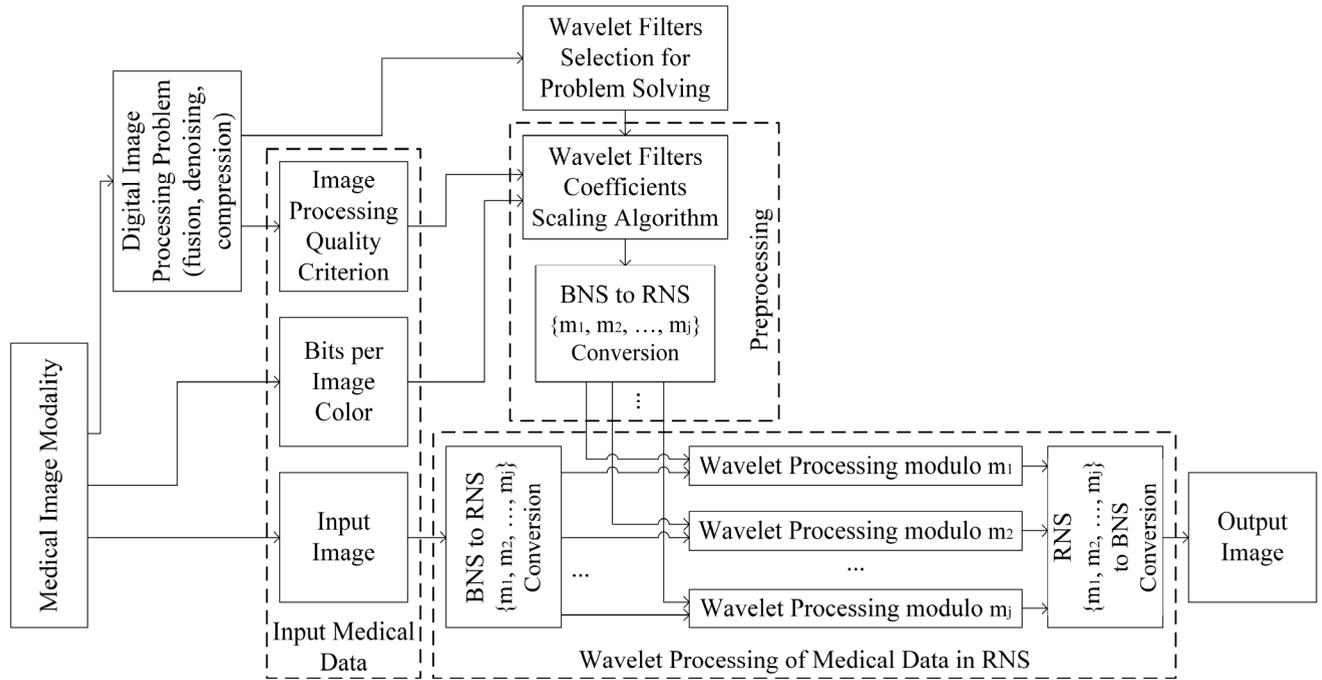


FIGURE 3. The proposed approach to wavelet processing of medical data in the residue number system.

in subsection IIIA. RNS-based calculations and hardware architecture features are represented in subsection IIIB.

### A. 3D IMAGES WAVELET PROCESSING METHOD AND ALGORITHM

WP is performed by 3D image processing using a filter bank which transforms the input data into time-frequency form, processing this data and performing the inverse transform, as shown in Fig. 4. A set of lowpass and highpass wavelet filters  $F$  for analysis (LD – lowpass decomposition; HD – highpass decomposition) and synthesis (LR – lowpass reconstruction; HR – highpass reconstruction) is a filter bank main component [68]. These filters consist of coefficients  $f_{F,i}$ , where  $i = 1, \dots, k$  is the index number and  $k$  is the number of coefficients. The greater the wavelet coefficients number, the better it separates the low-frequency and high-frequency signal components, but the higher the computational complexity. Wavelets with a large coefficient number are advisable when processing medical images with high color depth. Filter bank also includes downsampling  $\downarrow 2$ , upsampling  $\uparrow 2$ , and summation  $\oplus$  operators. We get 8 coefficients sets  $\{LLL, LLH, LHL, LHH, HLL, HLH, HHL, HHH\}$ . For example,  $LLH$  is a set of coefficients obtained from three consecutive overlays of wavelet filters on the image: lowpass  $L$ ; again lowpass  $L$ ; highpass  $H$ . They can be divided into approximating  $\{LLL\}$  (low-frequency component corresponding and image main information containing) and detailing  $\{LLH, LHL, LHH, HLL, HLH, HHL, HHH\}$  (high-frequency component corresponding and image detailed information containing) from the original image analysis according to the scheme in Fig. 4.

The main WP computational complexity is to use the repeated execution of a 3D image  $I(x, y, z)$  convolution with wavelet filters  $F$  as follows

$$\begin{aligned}
 I'(x, y, z) &= \sum_{i=1}^k I(x - i, y, z) \cdot f_{F,i}, \\
 I''(x, y, z) &= \sum_{i=1}^k I'(x, y - i, z) \cdot f_{F,i}, \\
 I'''(x, y, z) &= \sum_{i=1}^k I''(x, y, z - i) \cdot f_{F,i}, \quad (1)
 \end{aligned}$$

where:  $0 \leq x \leq X - 1$ ,  $0 \leq y \leq Y - 1$  and  $0 \leq z \leq Z - 1$  are the spatial coordinates of image  $I(x, y, z)$ ;  $I'(x, y, z)$ ,  $I''(x, y, z)$  and  $I'''(x, y, z)$  are the convolution results by rows, columns, and frames, respectively. The accuracy choice of wavelet filter coefficients representation is an important problem. This accuracy mainly affects the computational complexity of WP. Bit-width of type  $r = 2^\gamma$  ( $\gamma \in \mathbb{Z}$  and  $\gamma > 0$ ) is used to represent the wavelet filter coefficients for medical image WP in modern software packages due to the architecture peculiarities. Modern hardware accelerators have a more flexible architecture. The accuracy choice of wavelet filter coefficients representation at which low computational complexity will be achieved due to an insignificant loss in the 3D medical image WP quality is essential. We propose a new WP method using SFC. Bit-width  $r$  of SFC representation in the device memory primarily depends on the scaling factor (SF)  $n$ . Wavelet filter coefficients are converted to fixed-point

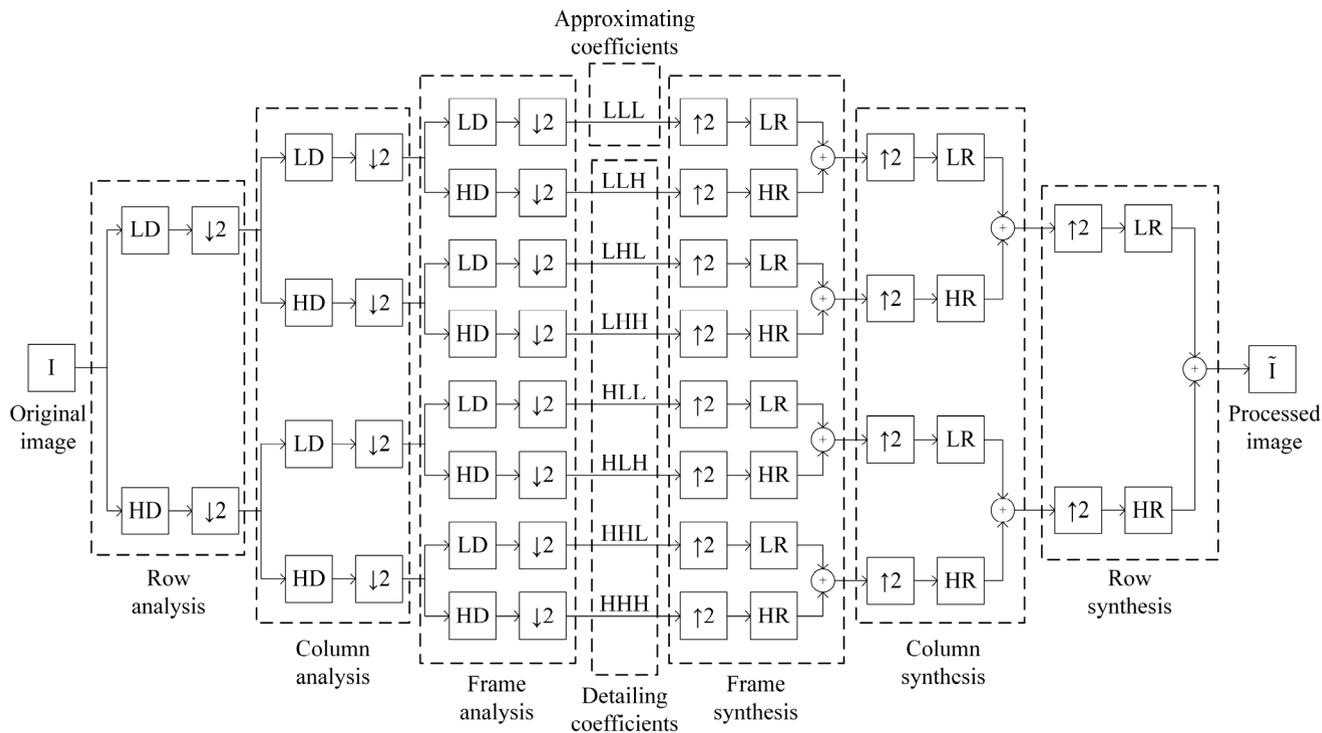


FIGURE 4. 3D medical image wavelet processing scheme.

format as follows

$$f_{F,i}^* = \lceil 2^n f_{F,i} \rceil. \tag{2}$$

Bit-width  $r$  of SFC calculated as follows

$$\begin{aligned} r &= \max_{1 \leq i \leq k} \{ \lceil \log_2 (| \lceil 2^n f_{F,i} \rceil | + 1) \rceil + 1 \} \\ &= \max_{1 \leq i \leq k} \{ \lceil \log_2 (| f_{F,i}^* | + 1) \rceil + 1 \}. \end{aligned} \tag{3}$$

Wavelet filters  $SLD, SHD, SLR, SHR$  consist of SFC  $f_{F,i}^*$ , and are used for 3D medical image  $I$  WP by the scheme in Fig. 5. For example,  $SLD$  is the  $LD$  (lowpass decomposition) filter with scaled coefficients. The convolutions and summations results normalization is performed as follows

$$\tilde{I} = \lfloor 2^{-6n} I^* \rfloor. \tag{4}$$

The proposed method has the following advantages.

1. The use of SFC with low SF reduces the computational complexity of 3D medical image WP and improves hardware device performance.
2. Integer calculations open up the possibility of time-efficient implementation on computing devices such as FPGAs using a fixed-point format for data storing and processing.
3. Integer calculations allow using RNS parallel computing to increase their speed.
4. Multiplication and division by  $2^n$  in the number binary notation corresponding to a comma shift by  $n$  bits to the right or left, respectively, significantly simplifies their realization.

5. Rounding errors have different signs and partially compensate for each other.

6. Rounding operations applying in a given order require fewer hardware resources than rounding operations to the nearest integer. Rounding up is performed beforehand.

All calculations performed by digital devices have an error due to the digital data format. Calculation error of WP occurs when wavelet filter coefficients are pre-scaling. It increases at each stage of 3D medical image WP according to the scheme in Fig. 5: when convolution with filters; when summing the convolution results; when normalizing the obtained values. The proposed method evaluates the 3D medical image WP quality as follows

$$PSNR = 10 \log_{10} \left( M^2 / MSE \right), \tag{5}$$

where:  $M$  is the maximum image voxels brightness value;  $MSE$  is the mean squared error, depends on SF and calculating by formula (A.1). The processing quality is considered high if  $PSNR \geq Q$  where  $Q$  is the medical image processing high-quality threshold calculating as follows

$$Q = C \cdot \log_2 (M + 1), \tag{6}$$

where  $C$  is the image processing quality factor. A threshold  $Q = 40$  dB is used for images with 8 bits per color (BPC) and  $M = 255$ . He describes the difference between 2D [69] or 3D [70] images that are not visible to the human eye. The factor  $C = \log_2(255 + 1)/40 = 5$  for images with  $BPC = 8$  by (6). Thus,  $Q = 5 \log_2(4095 + 1) = 60$  [8] and

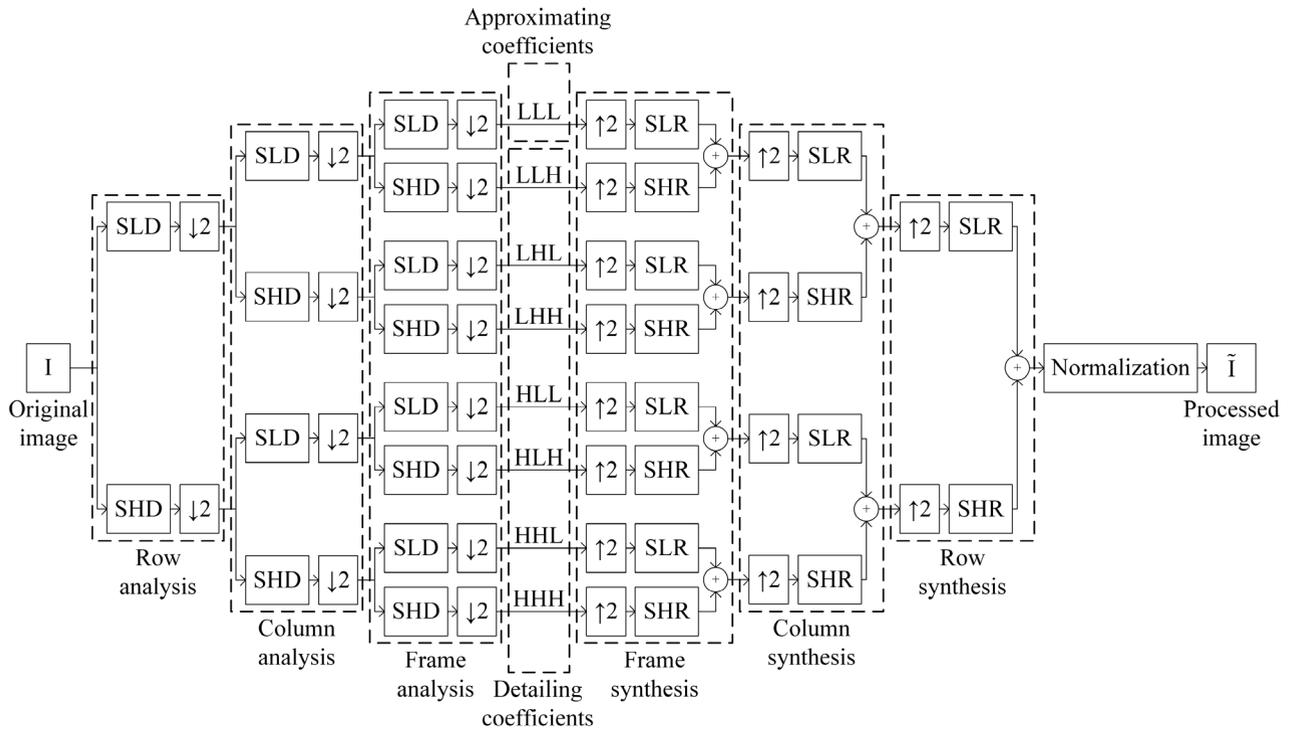


FIGURE 5. The proposed method scheme of 3D medical image wavelet processing using scaled filter coefficients.

$Q = 5 \log_2(65535 + 1) = 80$  for images with  $BPC = 12$  and  $BPC = 16$  respectively. Image parameters ( $BPC$  and  $M$ ), the wavelet coefficients number  $k$ , and the accuracy parameters of the SFC representation in the device memory (SF  $n$  and SFC bit-width  $r$ ) significantly affect the 3D medical image WP. The wavelet type has a minor effect. Image parameters are unchanged and depend on the medical image processing system properties. The wavelet type and a specific wavelet with the number  $k$  of the coefficients are selected depending on the processed image features and the conditions of a particular problem being solved. Thus, the accuracy of the SFC representation is the last important question that needs to be answered for high-performance and high-quality 3D medical image WP implementation on modern hardware accelerators.

The proposed Algorithm 1 calculates SFC  $f_{F,i}^*$ , and their bit-width  $r$  with known quality criteria and parameters of both input image and wavelet used. This algorithm is based on the minimum SF  $n$  calculation that satisfies the established image processing high-quality threshold  $Q$ . The initial SF is calculated using an approximate formula and then changes depending on the error. Wavelet filter coefficients scaling is performed after the minimum SF calculates at which 3D medical image has a high quality according to (6) after WP with SFC.

The known SFC and their bit-width open up the possibility of an efficient hardware implementation of the 3D medical image WP method in fixed-point format. The integer

calculations allow using RNS parallel computing to increase their speed. RNS-based calculations are described further.

**B. HARDWARE ACCELERATOR FEATURES WITH PARALLELIZED CALCULATIONS IN THE RESIDUE NUMBER SYSTEM FOR MEDICAL IMAGE PROCESSING**

RNS offers several advantages over the traditional positional number system [39]. The representation of high bitness numbers as low bitness residues is their main. Thus, the computational complexity is reduced for each calculation channel. Furthermore, modulo computations are performed independently, and there is no propagation of carrying between the channels. Thus, an error in one calculation channel does not propagate to others. The operations of sign determining, numbers comparing, and division called non-modular are difficult to implement in RNS and require the number positional characteristic calculation. Therefore, calculations in RNS are used to solve problems with a predominant amount of addition, subtraction, and multiplication operations. The digital filtering underlying WP is one such problem.

The RNS moduli type affects the computational speed, so the moduli selection is a significant task in RNS-based application devices. Moduli set must provide a sufficient dynamic range to represent numbers in RNS uniquely. Moduli must be balanced so that the operations execution time for each channel is approximately the same. RNS moduli of particular types  $2^\alpha$  and  $2^\alpha \pm 1$  ( $\alpha \in N$  where  $N$  is the

**Algorithm 1** Wavelet Filter Coefficients Scaling

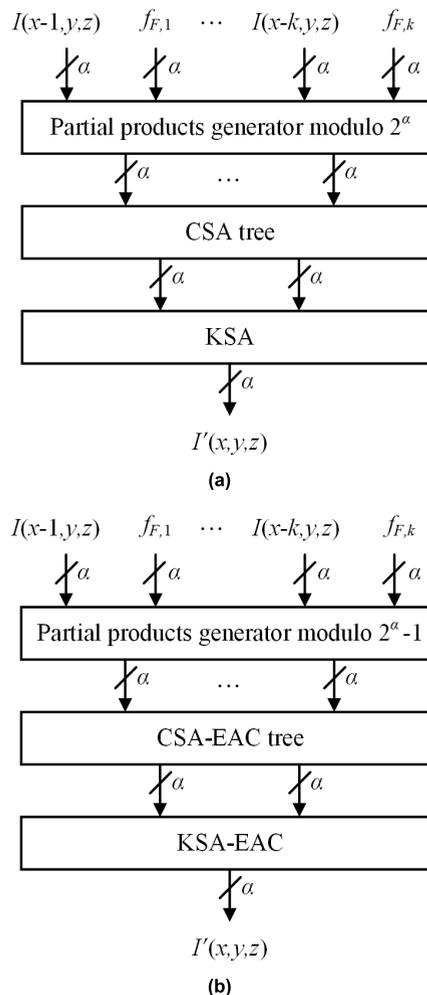
**Input:** Maximum image voxels brightness value  $M$ , image processing quality factor  $C$ , filter coefficients number  $k$ , filter coefficients  $f_{F,i}$

**Output:** SFC  $f_{F,i}^*$ , SFC bit-width  $r$

- 1:  $n \leftarrow \log_2(M + 1) + C - 1$  Calculate the initial SF
- 2:  $Q \leftarrow C \cdot \log_2(M + 1)$  Calculate the image processing high-quality threshold by (6)
- 3:  $PSNR \leftarrow 10\log_{10}(M^2/MSE)$  Calculate the image processing quality by (5) where  $MSE$  depends on SF and calculated by formula (A.1)
- 4: **if**  $PSNR \geq Q$  **then**
- 5:     **do**
- 6:          $n \leftarrow n - 1$
- 7:          $PSNR \leftarrow 10\log_{10}(M^2/MSE)$
- 8:         **while**  $PSNR \geq Q$
- 9:          $n \leftarrow n + 1$
- 10:     **else**
- 11:         **do**
- 12:              $n \leftarrow n + 1$
- 13:              $PSNR \leftarrow 10\log_{10}(M^2/MSE)$
- 14:             **while**  $PSNR < Q$
- 15:         **end if**
- 16:  $f_{F,i}^* \leftarrow \lceil 2^n f_{F,i} \rceil$  Calculate SFC by (2)
- 17:  $r = \max_{1 \leq i \leq k} \left\{ \lceil \log_2(|f_{F,i}^*| + 1) \rceil + 1 \right\}$  Calculate SFC bit-width by (3)

natural numbers) are hardware-friendly and allow to avoid the resource-intensive modulo division operation. The design of arithmetic devices modulo type  $2^\alpha$  is reduced to implementing the  $\alpha$ -bit devices in BNS. The moduli pairwise coprime requirement allows us to use only one modulus of type  $2^\alpha$  since the other moduli must be odd. Moduli of types  $2^\alpha$  and  $2^\alpha \pm 1$  enable efficient modulo summing techniques are similar to those used in BNS [71]. The modulo  $2^\alpha + 1$  calculation channel requires introducing additional logic according to the “diminished-1” method to track the zero codewords. It is undesirable for the system with minimal hardware resources and processing time [72]. Thus, type  $2^\alpha - 1$  moduli are preferable for use as odd RNS moduli.

3D medical image WP is performed using the operations of addition and multiplication by constants according to (1). Partial products can be precomputed using the compression technique [73] since the constants are known as apriori. Next, partial product addition is performed using a multi operand modular adder (MOMA) which consists of a carry-save adder (CSA) tree [89] and Kogge-Stone adder (KSA) [74].  $\alpha$ -bit adders are used for modulo  $2^\alpha$  calculations.  $\alpha$ -bit adders with end-around-carry (EAC) technique [71] are used for modulo  $2^\alpha - 1$  calculations. According to (1), the convolution device scheme is shown in Fig. 6. The  $\alpha$ -bit adders bit-width shown in Fig. 6 is significantly less than the full dynamic range bit-width. Thus, MOMAs reduce the arithmetic operations computational complexity and increase the RNS-based device speed. Experiments on software and hardware implementation of the developed method and the proposed algorithm for 3D medical image WP are presented further.



**FIGURE 6.** The convolution device scheme for 3D image  $I(x, y, z)$  and wavelet filter coefficients  $f_{F,i}$  using CSA tree and KSA: (a) modulo  $2^\alpha$ ; (b) modulo  $2^\alpha - 1$  with EAC technique.

**IV. EXPERIMENTS**

This section is divided into 2 subsections. Subsection IVA presents the software implementation of 3D computed tomogram WP with SFC using various scaling parameters and its comparison with the proposed algorithm. Subsection IVB describes the hardware implementations of WP for 3D medical images with SFC and parallelized calculations in RNS and their comparison with state-of-the-art solutions.

**A. SOFTWARE IMPLEMENTATION OF MEDICAL IMAGE WP WITH SFC**

Medical image WP with SFC was implemented in the MATLAB software version R2020b for the 3D computed tomogram of the paranasal sinuses “CT\_par\_sin” is the grayscale 12-bit image of size  $390 \times 390 \times 390$ . Filters coefficients  $f_{F,i}$  of the Daubechies wavelets  $db(k/2)(k = 2, 4, 6, \dots, 20)$  were obtained, scaled using different  $2^n$  by (2), and converted to the fixed-point format. Computed tomogram WP with SFC implemented using (1) and (4) according to the scheme in Fig. 5. An example of a 3D

tomographic image “CT\_par\_sin” WP is shown in Fig. 7. The results ( $PSNR$ , dB) of 3D grayscale 12-bit computed tomogram “CT\_par\_sin” software WP using Daubechies wavelets  $db(k/2)$  with the number  $k = 2, 8, 14, 20$  of the coefficients scaled by  $n = 1, 2, 3, \dots, 17$  is shown in Fig. 8. The minimum SFs for high-quality according to (6) WP with SFC of 3D grayscale 12-bit computed tomogram “CT\_par\_sin” using Daubechies wavelets  $db(k/2)$  with the number  $k = 2, 8, 14, 20$  of the coefficients is shown in Fig. 9. Hardware implementation of 3D medical image WP with different color bit-depth using various SFC with bit-width  $r$  on hardware FPGA accelerators is presented further.

### B. HARDWARE IMPLEMENTATIONS OF MEDICAL IMAGE WP WITH SFC AND PARALLELIZED CALCULATIONS IN RNS

Medical image WP with SFC was implemented in the Xilinx Vivado 2020.2 for the target device Kintex UltraScale xcku115-flvf-1924-3-e. The run strategy is “Flow\_PerfOptimized\_high.” The higher the image color depth, the greater the calculation accuracy needed for high-quality processing, and the more bits must be allocated for wavelet filter coefficients. Therefore, different wavelets are selected to process 3D medical images with varying depths of color. Hardware implementation was carried out for WP with SFC using Daubechies wavelet  $db2$  with 4 coefficients, symlets  $sym6$  with 12 coefficients, and coiflet  $coif4$  with 24 coefficients to process images with  $BPC = 8, 12, 16$ , respectively. Filters coefficients  $f_{F,i}$  were scaled using different  $2^n$  by (2) and converted to fixed-point format with bit-widths  $r$  calculated by (3). Various SFs and bit-widths  $r$  corresponding to the known and proposed implementation methods are used and are presented in Table 1. Hardware implementation results of 3D medical image WP with SFC by various state-of-the-art and proposed methods from Table 1 are shown in Table 2.

Implementation of medical image WP with SFC and parallelized calculations in RNS was carried out using Daubechies wavelet  $db2$ . Various SFs and bit-widths  $r$  corresponding to the known and proposed implementation methods presented in Table 1 were used. Both the RNS dynamic range

requirements [76] were calculated, and the moduli sets were selected for each method of 8-bit image processing from Table 1 (Table 3). Hardware implementation results of 3D medical image WP using different filter coefficients bit-width  $r$  and RNS parallel computing and their comparison with BNS from are presented in Table 4.

A description and discussion of all the software and hardware implementation results obtained are presented further.

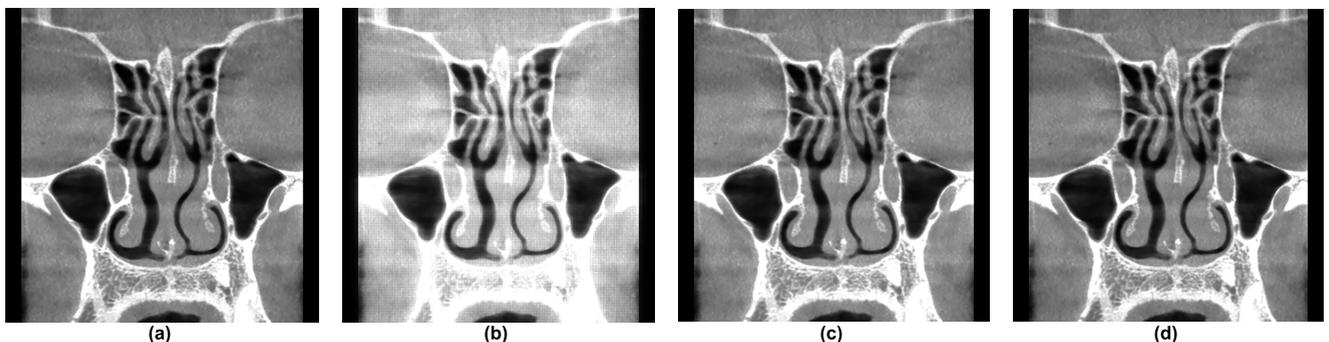
## V. RESULTS AND DISCUSSION

This section is divided into 2 subsections corresponding to the subsections with software and hardware implementations from Section IV.

### A. SOFTWARE IMPLEMENTATION RESULTS OF MEDICAL IMAGE WP WITH SFC

Fig. 7 shows that the higher the SFC representation accuracy, the lower the calculation error and its influence on the 3D computed tomogram processing quality. The computed tomogram in Fig. 7b processed using low precision SFC ( $n = 6$ ) has a low-quality ( $PSNR = 18.18$  dB). It is heavily lighted. Image distortions appear as lighting since the calculation errors for each voxel have the same sign and are excessive according to (2). WP calculation error leads to the distribution of distortion over the entire image since each filter is used to process all image voxels. The image seems to be divided into squares at low processing quality. Upsampling leads to this phenomenon, short brightness leaps, and the visible boundaries appearance between fragments of size  $2 \times 2$ . These fragments will have a size  $2 \times 2 \times 2$  in a full 3D tomogram representation. This regularity can be used to image quality assessment in lossy compression since similar distortions appear [21], [32].

The computed tomogram in Fig. 7c processed using average precision SFC ( $n = 9$ ) has an average quality ( $PSNR = 35.71$  dB). Distortions are less pronounced than in Fig. 7b and are most noticeable in image fragments with bone tissue since their brightness is higher than in the other image fragments. The higher the brightness, the greater the WP calculation error, and the more noticeable the image processed distortions. This regularity is important since medical diagnosis



**FIGURE 7.** Example of 3D tomogram “CT\_par\_sin” (120-th coronal frame) wavelet processing with SFC using Daubechies wavelet  $db7$  and various SF  $n$ : (a) original image; processed image: (b)  $n = 6$ ,  $PSNR = 18.18$  dB; (c)  $n = 9$ ,  $PSNR = 35.71$  dB; (d)  $n = 13$ ,  $PSNR = 62.05$  dB.

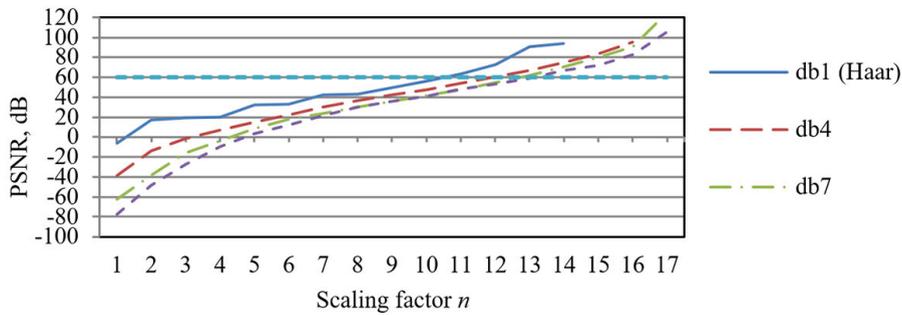


FIGURE 8. Results of 3D grayscale 12-bit computed tomogram “CT\_par\_sin” wavelet processing with SFC using Daubechies wavelets.

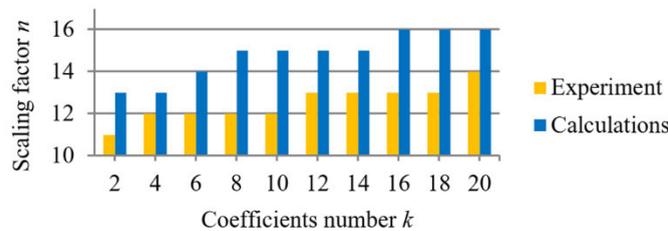


FIGURE 9. The minimum SF n at which the results of wavelet processing with SFC of 3D grayscale 12-bit computed tomogram “CT\_par\_sin” using Daubechies wavelets achieves high quality.

TABLE 1. Scaling factor values n used for hardware implementation.

Method	Image BPC	Wavelet	Scaling factor n	Coefficients bit-width r
IEEE Standard 1076-2019 [75]	8	db2	31	32
	12	sym 6	31	32
	16	coif 4	31	32
Method [38]	8	db2	11	12
Method [54]	8	db2	15	16
Proposed	8	db2	10	11
	12	sym 6	15	16
	16	coif 4	19	20

TABLE 2. Hardware implementation results of 3D medical image wavelet processing with scaled filter coefficients.

Method	Image BPC	Coefficients bit-width r	Performance, megavoxels/s	Hardware resources, LUTs
IEEE Standard 1076-2019 [75]	8	32	58.133	4094
	12	32	46.206	11989
	16	32	33.464	24325
Method [38]	8	12	72.417	1465
Method [54]	8	16	66.631	1760
Proposed	8	11	<b>82.488</b>	<b>1250</b>
	12	16	46.703	4527
	16	20	42.084	11749

uses the brightest image fragments mainly. CT is used to gain knowledge of high-density tissues primarily. Human

TABLE 3. Dynamic range requirements and moduli sets of residue number system for hardware implementation.

Method	Coefficients bit-width r	Dynamic range D	Moduli set {m <sub>1</sub> , m <sub>2</sub> , ..., m <sub>j</sub> }
Method [38]	12	1612110	{2 <sup>8</sup> - 1, 2 <sup>9</sup> - 1, 2 <sup>9</sup> }
Method [54]	16	25796820	{2 <sup>9</sup> - 1, 2 <sup>10</sup> - 1, 2 <sup>10</sup> }
IEEE Standard 1076-2019 [75]	32	1690601720340	{2 <sup>14</sup> - 1, 2 <sup>15</sup> - 1, 2 <sup>15</sup> }
Proposed	11	806820	{2 <sup>7</sup> - 1, 2 <sup>8</sup> - 1, 2 <sup>8</sup> }

TABLE 4. Hardware implementation results of 3D medical image wavelet processing with SFC in RNS and BNS.

Method	Coefficients bit-width r	Performance, megavoxels/s		Hardware resources, LUTs	
		RNS	BNS	RNS	BNS
Method [38]	12	182.282	72.417	5715	1465
Method [54]	16	155.063	66.631	6654	1760
IEEE Standard 1076-2019 [75]	32	103.359	58.133	12731	4094
Proposed	11	<b>208.986</b>	<b>82.488</b>	<b>4820</b>	<b>1250</b>

bones are this tissue in this case. Distortions introduced into fragments of the “CT\_par\_sin” image with bone tissue are the most noticeable and significant. Although most images do not contain visible distortions, the most important fragments are distorted enough to influence the diagnosis.

The computed tomogram in Fig. 7d processed using high precision SFC (n = 13) has a high-quality (PSNR = 62.05 > Q = 60 dB) according to (6). This image does not contain

visible distortions, indistinguishable by eye from the image in Fig. 7a, and can be used for medical diagnosis. Nevertheless, the images obtained using different modalities are very diverse. They have many parameters such as size and color depth and features related to the imaging modality, the organs being diagnosed, both properties and settings of the diagnostic device, and others. A general approach to 3D medical image WP quality evaluating is proposed in this paper. However, the expert can determine the image processing high-quality threshold based on the conditions and circumstances of the clinical diagnosis in each case. The developed WP method with SFC and the proposed coefficient scaling algorithm can be used in any approach to assessing the image processing quality using the PSNR criterion. The diagnosis can be carried out by both a medical expert and a machine. Thus, the proposed approach shown in Fig. 3 can be used in artificial intelligence systems using DWT-based convolutional neural networks for various medical image computing problem solving such as segmentation, classification, and pattern recognition [1].

Experimental results of 3D grayscale 12-bit computed tomogram “CT\_par\_sin” WP presented in Fig. 8 and 9 confirm the theoretical analysis results correctness from Section 3. The more coefficients the wavelet filters contain, the higher the accuracy of their representation in the device memory for high-quality medical image processing should be. Practical results have shown that filter coefficients can be scaled 1-3 bits less than the theoretical analysis requires. The theoretical analysis orientation explains this difference to the worst-case identifying at which the calculation error has the maximum impact on the image processing quality. A probabilistic approach proposed in [30] based on statistical methods can be used for the non-redundant SF calculation and further computational complexity reduction. However, using the SFC representation accuracy according to the theoretical analysis results guarantees high-quality 3D medical image WP. WP-based methods development for medical image fusion [5], denoising [9], and compression [22] based on the proposed approach is one of the further research promising areas. Hardware implementation results and discussion of 3D medical image WP with different color bit-depth using various SFC with bit-width  $r$  on hardware FPGA accelerators are presented further.

### B. HARDWARE IMPLEMENTATIONS RESULTS OF MEDICAL IMAGE WP WITH SFC AND PARALLELIZED CALCULATIONS IN RNS

The higher the device performance, the better, as it can process more information in the same amount of time. The lower the hardware cost, the better. Lower hardware costs result in lower power consumption and smaller device size requirements. Hardware implementation results from Table 2 shows that:

1. All considered methods allow achieving high-quality image processing. The quality of image processing using different methods is approximately equal.

2. The use of the proposed approach (shown in Fig. 3) for 3D 8-bit medical image WP with SFC: increased the device performance by 13.91% compared to [38], by 23.80% compared to [54], and by 41.90% compared to [75]; reduced hardware resources by 14.68% compared to [38], 28.98% compared to [54], and 69.47% compared to [75].

3. The proposed approach for 3D 12-bit medical image WP increased the device performance by 1.08% and reduced the hardware resources by 62.24% compared to [75].

4. The proposed approach for 3D 16-bit medical image WP increased the device performance by 25.76% and reduced the hardware resources by 51.70% compared to [75].

Thus, hardware implementation results of 3D medical images WP using various color depths, proposed and known state-of-the-art methods (Table 1) confirm that the SFC representation accuracy correlates with the WP computational complexity. Which directly affects the time, hardware, and energy costs of medical image processing devices. Tomographic image WP with SFC does not lead to the visible distortion appearance simultaneously. We conclude that the proposed approach makes it possible to increase the device performance and reduce the hardware resources while maintaining high-quality WP of 3D medical images with different color depths. Compared with the known methods in the hardware resources, the proposed approach significantly exceeds the device performance gain in percentage terms in each case. SFC-based WP using the lifting scheme further reduces hardware resources for low-area device priority applications. The lifting scheme called the second generation wavelet transform has lower computational complexity than the classic wavelet transforms but has disadvantages [22], [41]. Sequential data processing in the lifting scheme increases the arithmetic operations delay and decreases the transformation speed compared to parallel data processing in the classic wavelet transform. We assume that the lifting scheme will mostly negate the performance gain of the medical image WP device obtained using the proposed approach but will lead to an even more significant hardware resources reduction.

Hardware implementation results of 3D 8-bit medical image WP are presented in Table 4 using proposed and state-of-the-art methods (Table 3) and show that:

1. Parallel 3-modulo RNS calculations increased the device performance by 1.78-2.53 times compared to using BNS, increasing the hardware resources by 3.11-3.90 times.

2. Both SFC-based WP and RNS parallel computing combinations increased the device performance by 2.89-3.59 times, increasing the hardware resources by 1.18-3.29 times.

Moduli sets of a particular type made it possible to use effective summation in CSA and KSA blocks with the EAC technique, which led to a significant calculation speed increase. However, RNS to BNS conversion operation is very resource-intensive, partially neutralizes the device performance gain, and greatly affects the hardware resources. This approach is advisable for medical imaging applications with performance-priority devices [43]. The developed FPGA accelerators are cheap, practical, and can be introduced into

the mass market. However, the proposed approach implementation on FPGA does not fully realize the RNS arithmetic potential. FPGAs contain a finite set of elements, and connections between them are programmed only. The developed method and the proposed algorithm implementation on ASIC devices allow the significantly higher performance improvement of 3D medical image WP with a considerably smaller hardware resources increase. ASIC has more fine-tuning possibility for a specific task is one of the further research promising directions [35].

## VI. CONCLUSION

In this paper, we proposed an approach to the 3D medical image WP device performance improvements based on SFC using and RNS parallel computing. The developed SFC-based WP method has reduced the computational complexity. The proposed method of the 3D image WP accuracy evaluation estimated the SF influence on the medical image processing quality. Proven that SFC for 3D medical image high-quality WP with reduced computational complexity can be calculated using the proposed algorithm. Software implementation of the proposed approach has achieved high-quality WP of 3D 12-bit computed tomogram. Hardware implementation of 3D medical image WP using the proposed FPGA accelerators and comparison with state-of-the-art solutions have shown: the use of SFC-based WP increases the device performance up to 41.90% and reduces the hardware resources up to 69.47%; the use of RNS parallel computing increases the device performance by 1.78-2.53 times compared to using BNS with an increase in the hardware resources by 3.11-3.90 times; the both SFC-based WP and RNS parallel computing combination increases the device performance by 2.89-3.59 times with an increase in the hardware resources by 1.18-3.29 times. The improvement of device performance is achieved while maintaining high-quality medical image WP in PSNR.

In further work, we plan to apply the proposed approach for: performance improvement of the non-medical image processing systems; DWT-based neural networks developing for medical image segmentation, classification, and pattern recognition; developed methods modification by a probabilistic approach using based on statistical methods to further computational complexity reducing; WP-based methods development for medical image fusion, denoising, and compression; wavelet methods development based on the lifting scheme to the hardware resources further reduce; developed methods implementation on ASIC to the device performance further improve.

## APPENDIX

Calculation error of WP occurs when wavelet filter coefficients are pre-scaling. It increases at each stage of 3D medical image WP according to the scheme in Figure 5: when convolution with filters; when summing the convolution results; when normalizing the obtained values. Analyzing and synthesizing filters contain the same coefficients set according

to formula (1). Thus, the errors of filters *SLD* and *SLR* will be equal and the errors of filters *SHD* and *SHR*. Therefore, filters *SLD* and *SLR* are denoted as *L*, and filters *SHD* and *SHR* as *H*. The coefficient values will be divided into several groups when calculations are performed. Even and odd filter coefficients, for example. The symbol  $\lambda$  is added to denote a group, and it means that calculations are carried out for  $\lambda$  coefficients groups.

*Stage 1. Wavelet Filter Coefficients Scaling (Preprocessing):*

$$T_1^{[L]} = \sum_{i=1}^k 2^n f_{L,i} = 2^n \sum_{i=1}^k f_{L,i} = 2^n \cdot \sqrt{2} = 2^{n+1/2},$$

where  $T_1^{[L]}$  is the exact value of the filter *L* coefficients sum;

$$T_1^{[H]} = \sum_{i=1}^k 2^n f_{H,i} = 2^n \sum_{i=1}^k f_{H,i} = 2^n \cdot 0 = 0,$$

$$T_{1,1}^{[L]} = \sum_{i=1}^{k/2} 2^n f_{L,2i-1};$$

$$T_{1,2}^{[L]} = \sum_{i=1}^{k/2} 2^n f_{L,2i}, \quad T_{1,1}^{[H]} = \sum_{i=1}^{k/2} 2^n f_{H,2i-1},$$

$$T_{1,2}^{[H]} = \sum_{i=1}^{k/2} 2^n f_{H,2i};$$

$$E_1^{[L]} = \sum_{i=1}^k ( \lceil 2^n f_{L,i} \rceil - 2^n f_{L,i} ),$$

where  $E_1^{[L]}$  is the absolute rounding errors sum of filter *L* coefficients;

$$E_1^{[H]} = \sum_{i=1}^k ( \lceil 2^n f_{H,i} \rceil - 2^n f_{H,i} );$$

$$E_{1,1}^{[L]} = \sum_{i=1}^{k/2} ( \lceil 2^n f_{L,2i-1} \rceil - 2^n f_{L,2i-1} );$$

$$E_{1,2}^{[L]} = \sum_{i=1}^{k/2} ( \lceil 2^n f_{L,2i} \rceil - 2^n f_{L,2i} );$$

$$E_{1,1}^{[H]} = \sum_{i=1}^{k/2} ( \lceil 2^n f_{H,2i-1} \rceil - 2^n f_{H,2i-1} );$$

$$E_{1,2}^{[H]} = \sum_{i=1}^{k/2} ( \lceil 2^n f_{H,2i} \rceil - 2^n f_{H,2i} ).$$

*Stage 2. Rows Convolution (Analysis):*

$$T_2^{[L]} = T_1^{[L]} \cdot M,$$

where:  $T_2^{[L]}$  is the calculations exact value at the 2nd stage obtained after image convolution by filter *L*; *M* is the maximum image voxels brightness value;

$$E_2^{[L]} = E_1^{[L]} \cdot M,$$

where  $E_2^{[L]}$  is the maximum absolute error at the 2nd stage obtained after image convolution by filter  $L$ ;

$$E_2^{[H]} = E_1^{[H]} \cdot M.$$

All  $T_l^{[F]} = 0$  ( $l > 1$ ), where  $F$  an image convolution filter sequence containing  $H$  since  $T_1^{[H]} = 0$ .

*Stage 3. Columns Convolution (Analysis):*

$$T_3^{[LL]} = T_2^{[LL]} \cdot T_1^{[L]},$$

where  $T_3^{[LL]}$  is the calculations exact value at the 3rd stage obtained after double sequential image convolution by  $L$ ;

$$E_3^{[LL]} = \left(T_2^{[L]} + E_2^{[L]}\right) \left(T_1^{[L]} + E_1^{[L]}\right) - T_3^{[LL]},$$

where  $E_3^{[LL]}$  is the maximum absolute error at the 3rd stage obtained after double sequential image convolution by  $L$ ;

$$E_3^{[LH]} = \left(T_2^{[L]} + E_2^{[L]}\right) E_1^{[H]}, \quad E_3^{[HL]} = E_2^{[H]} \left(T_1^{[L]} + E_1^{[L]}\right);$$

$$E_3^{[HH]} = E_2^{[H]} \cdot E_1^{[H]}.$$

*Stage 4. Frames Convolution (Analysis):*

$$T_4^{[LLL]} = T_3^{[LL]} \cdot T_1^{[L]}, \quad E_4^{[LLL]} = \left(T_3^{[LL]} + E_3^{[LL]}\right) \left(T_1^{[L]} + E_1^{[L]}\right) - T_4^{[LLL]},$$

$$E_4^{[LLH]} = \left(T_3^{[LL]} + E_3^{[LL]}\right) E_1^{[H]},$$

$$E_4^{[LHL]} = E_3^{[LH]} \left(T_1^{[L]} + E_1^{[L]}\right);$$

$$E_4^{[LHH]} = E_3^{[LH]} \cdot E_1^{[H]}, \quad E_4^{[HLL]} = E_3^{[HL]} \left(T_1^{[L]} + E_1^{[L]}\right);$$

$$E_4^{[HLH]} = E_3^{[HL]} \cdot E_1^{[H]}, \quad E_4^{[HHL]} = E_3^{[HH]} \left(T_1^{[L]} + E_1^{[L]}\right);$$

$$E_4^{[HHH]} = E_3^{[HH]} \cdot E_1^{[H]}.$$

*Stage 5. Frames Convolution (Synthesis):* ( $\lambda = 1, 2$ )

$$T_{5,\lambda}^{[LLL]} = T_4^{[LLL]} \cdot T_{1,\lambda}^{[L]},$$

$$E_{5,\lambda}^{[LLL]} = \left(T_4^{[LLL]} + E_4^{[LLL]}\right) \left(T_{1,\lambda}^{[L]} + E_{1,\lambda}^{[L]}\right) - T_{5,\lambda}^{[LLL]},$$

$$E_{5,\lambda}^{[LLHH]} = E_4^{[LLH]} \left(T_{1,\lambda}^{[H]} + E_{1,\lambda}^{[H]}\right),$$

$$E_{5,\lambda}^{[LHLL]} = E_4^{[LHL]} \left(T_{1,\lambda}^{[L]} + E_{1,\lambda}^{[L]}\right);$$

$$E_{5,\lambda}^{[LHHH]} = E_4^{[LHH]} \left(T_{1,\lambda}^{[H]} + E_{1,\lambda}^{[H]}\right),$$

$$E_{5,\lambda}^{[HLLL]} = E_4^{[HLL]} \left(T_{1,\lambda}^{[L]} + E_{1,\lambda}^{[L]}\right);$$

$$E_{5,\lambda}^{[HLHH]} = E_4^{[HLH]} \left(T_{1,\lambda}^{[H]} + E_{1,\lambda}^{[H]}\right),$$

$$E_{5,\lambda}^{[HHLL]} = E_4^{[HLL]} \left(T_{1,\lambda}^{[L]} + E_{1,\lambda}^{[L]}\right);$$

$$E_{5,\lambda}^{[HHHH]} = E_4^{[HHH]} \left(T_{1,\lambda}^{[H]} + E_{1,\lambda}^{[H]}\right).$$

*Stage 6. Frames Summation:* ( $\lambda = 1, 2$ )

$$E_{6,\lambda}^{[LL]} = E_{5,\lambda}^{[LLL]} + E_{5,\lambda}^{[LLHH]}, \quad E_{6,\lambda}^{[H]} = E_{5,\lambda}^{[LHLL]} + E_{5,\lambda}^{[LHHH]},$$

$$E_{6,\lambda}^{[HL]} = E_{5,\lambda}^{[HLLL]} + E_{5,\lambda}^{[HLHH]}, \quad E_{6,\lambda}^{[HH]} = E_{5,\lambda}^{[HHLL]} + E_{5,\lambda}^{[HHHH]}.$$

*Stage 7. Columns Convolution (Synthesis):*

$$T_{7,1}^{[LL]} = T_{5,1}^{[LLL]} \cdot T_{1,1}^{[L]}, \quad T_{7,2}^{[LL]} = T_{5,2}^{[LLL]} \cdot T_{1,1}^{[L]},$$

$$T_{7,3}^{[LL]} = T_{5,1}^{[LLL]} \cdot T_{1,2}^{[L]},$$

$$T_{7,4}^{[LL]} = T_{5,1}^{[LLL]} \cdot T_{1,2}^{[L]},$$

$$E_{7,1}^{[LL]} = \left(T_{5,1}^{[LLL]} + E_{6,1}^{[LL]}\right) \left(T_{1,1}^{[L]} + E_{1,1}^{[L]}\right) - T_{7,1}^{[LL]},$$

$$E_{7,2}^{[LL]} = \left(T_{5,2}^{[LLL]} + E_{6,2}^{[LL]}\right) \left(T_{1,1}^{[L]} + E_{1,1}^{[L]}\right) - T_{7,2}^{[LL]},$$

$$E_{7,3}^{[LL]} = \left(T_{5,1}^{[LLL]} + E_{6,1}^{[LL]}\right) \left(T_{1,2}^{[L]} + E_{1,2}^{[L]}\right) - T_{7,3}^{[LL]},$$

$$E_{7,4}^{[LL]} = \left(T_{5,2}^{[LLL]} + E_{6,2}^{[LL]}\right) \left(T_{1,2}^{[L]} + E_{1,2}^{[L]}\right) - T_{7,4}^{[LL]},$$

$$E_{7,1}^{[LH]} = E_{6,1}^{[LH]} \left(T_{1,1}^{[H]} + E_{1,1}^{[H]}\right),$$

$$E_{7,2}^{[LH]} = E_{6,2}^{[LH]} \left(T_{1,1}^{[H]} + E_{1,1}^{[H]}\right);$$

$$E_{7,3}^{[LH]} = E_{6,1}^{[LH]} \left(T_{1,2}^{[H]} + E_{1,2}^{[H]}\right),$$

$$E_{7,4}^{[LH]} = E_{6,2}^{[LH]} \left(T_{1,2}^{[H]} + E_{1,2}^{[H]}\right);$$

$$E_{7,1}^{[HL]} = E_{6,1}^{[HL]} \left(T_{1,1}^{[L]} + E_{1,1}^{[L]}\right),$$

$$E_{7,2}^{[HL]} = E_{6,2}^{[HL]} \left(T_{1,1}^{[L]} + E_{1,1}^{[L]}\right);$$

$$E_{7,3}^{[HL]} = E_{6,1}^{[HL]} \left(T_{1,2}^{[L]} + E_{1,2}^{[L]}\right),$$

$$E_{7,4}^{[HL]} = E_{6,2}^{[HL]} \left(T_{1,2}^{[L]} + E_{1,2}^{[L]}\right);$$

$$E_{7,1}^{[HH]} = E_{6,1}^{[HH]} \left(T_{1,1}^{[H]} + E_{1,1}^{[H]}\right),$$

$$E_{7,2}^{[HH]} = E_{6,2}^{[HH]} \left(T_{1,1}^{[H]} + E_{1,1}^{[H]}\right);$$

$$E_{7,3}^{[HH]} = E_{6,1}^{[HH]} \left(T_{1,2}^{[H]} + E_{1,2}^{[H]}\right),$$

$$E_{7,4}^{[HH]} = E_{6,2}^{[HH]} \left(T_{1,2}^{[H]} + E_{1,2}^{[H]}\right).$$

*Stage 8. Columns Summation:* ( $\lambda = \overline{1, 4}$ )

$$E_{8,\lambda}^{[L]} = E_{7,\lambda}^{[LL]} + E_{7,\lambda}^{[LH]}, \quad E_{8,\lambda}^{[H]} = E_{7,\lambda}^{[HL]} + E_{7,\lambda}^{[HH]}.$$

*Stage 9. Rows Convolution (Synthesis):*

$$T_{9,1} = T_{7,1}^{[LL]} \cdot T_{1,1}^{[L]}, \quad T_{9,2} = T_{7,2}^{[LL]} \cdot T_{1,1}^{[L]},$$

$$T_{9,3} = T_{7,3}^{[LL]} \cdot T_{1,1}^{[L]},$$

$$T_{9,4} = T_{7,4}^{[LL]} \cdot T_{1,1}^{[L]}, \quad T_{9,5} = T_{7,1}^{[LL]} \cdot T_{1,2}^{[L]},$$

$$T_{9,6} = T_{7,2}^{[LL]} \cdot T_{1,2}^{[L]},$$

$$T_{9,7} = T_{7,3}^{[LL]} \cdot T_{1,2}^{[L]}, \quad T_{9,8} = T_{7,4}^{[LL]} \cdot T_{1,2}^{[L]},$$

$$E_{9,1}^{[L]} = \left(T_{7,LL,1} + E_{8,1}^{[L]}\right) \left(T_{1,1}^{[L]} + E_{1,1}^{[L]}\right) - T_{9,1};$$

$$E_{9,2}^{[L]} = \left(T_{7,LL,2} + E_{8,2}^{[L]}\right) \left(T_{1,1}^{[L]} + E_{1,1}^{[L]}\right) - T_{9,2};$$

$$E_{9,3}^{[L]} = \left(T_{7,LL,3} + E_{8,3}^{[L]}\right) \left(T_{1,1}^{[L]} + E_{1,1}^{[L]}\right) - T_{9,3};$$

$$E_{9,4}^{[L]} = \left(T_{7,LL,4} + E_{8,4}^{[L]}\right) \left(T_{1,1}^{[L]} + E_{1,1}^{[L]}\right) - T_{9,4};$$

$$E_{9,5}^{[L]} = \left(T_{7,LL,1} + E_{8,1}^{[L]}\right) \left(T_{1,2}^{[L]} + E_{1,2}^{[L]}\right) - T_{9,5};$$

$$E_{9,6}^{[L]} = \left(T_{7,LL,2} + E_{8,2}^{[L]}\right) \left(T_{1,2}^{[L]} + E_{1,2}^{[L]}\right) - T_{9,6};$$

$$\begin{aligned}
E_{9,7}^{[L]} &= (T_{7,LL,3} + E_{8,3}^{[L]}) (T_{1,2}^{[L]} + E_{1,2}^{[L]}) - T_{9,7}; \\
E_{9,8}^{[L]} &= (T_{7,LL,4} + E_{8,4}^{[L]}) (T_{1,2}^{[L]} + E_{1,2}^{[L]}) - T_{9,8}; \\
E_{9,1}^{[H]} &= E_{8,1}^{[H]} (T_{1,1}^{[H]} + E_{1,1}^{[H]}), \quad E_{9,2}^{[H]} = E_{8,2}^{[H]} (T_{1,1}^{[H]} + E_{1,1}^{[H]}); \\
E_{9,3}^{[H]} &= E_{8,3}^{[H]} (T_{1,1}^{[H]} + E_{1,1}^{[H]}), \quad E_{9,4}^{[H]} = E_{8,4}^{[H]} (T_{1,1}^{[H]} + E_{1,1}^{[H]}); \\
E_{9,5}^{[H]} &= E_{8,1}^{[H]} (T_{1,2}^{[H]} + E_{1,2}^{[H]}), \quad E_{9,6}^{[H]} = E_{8,2}^{[H]} (T_{1,2}^{[H]} + E_{1,2}^{[H]}); \\
E_{9,7}^{[H]} &= E_{8,3}^{[H]} (T_{1,2}^{[H]} + E_{1,2}^{[H]}), \quad E_{9,8}^{[H]} = E_{8,4}^{[H]} (T_{1,2}^{[H]} + E_{1,2}^{[H]}).
\end{aligned}$$

Stage 10. Rows Summation: ( $\lambda = \overline{1,8}$ )

$$E_{10,\lambda} = E_{9,\lambda}^{[L]} + E_{9,\lambda}^{[H]}.$$

Stage 11. Normalization: ( $\lambda = \overline{1,8}$ )

$$E_{11,\lambda} = \left\lfloor 2^{-6n} E_{10,\lambda} \right\rfloor.$$

Values  $E_{11,\lambda}$  ( $\lambda = \overline{1,8}$ ) are the maximum resulting error of 3D medical image WP that calculate *MSE* by the formula

$$MSE = (1/8) \sum_{\lambda=1}^8 E_{11,\lambda}^2. \quad (A.1)$$

## ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for the valuable remarks made during the discussion of the work, which made it possible to improve the paper's quality significantly, and also would like to thank North-Caucasus Federal University for Supporting in the Contest of Projects Competition of Scientific Groups and Individual Scientists of North-Caucasus Federal University.

## REFERENCES

- [1] P. Kumar Mallick, S. H. Ryu, S. K. Satapathy, S. Mishra, G. N. Nguyen, and P. Tiwari, "Brain MRI image classification for cancer detection using deep wavelet autoencoder-based deep neural network," *IEEE Access*, vol. 7, pp. 46278–46287, 2019, doi: [10.1109/ACCESS.2019.2902252](https://doi.org/10.1109/ACCESS.2019.2902252).
- [2] W. Lv, S. Ashrafinia, J. Ma, L. Lu, and A. Rahmim, "Multi-level multi-modality fusion radiomics: Application to PET and CT imaging for prognostication of head and neck cancer," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 8, pp. 2268–2277, Aug. 2020, doi: [10.1109/JBHI.2019.2956354](https://doi.org/10.1109/JBHI.2019.2956354).
- [3] X. Wu, J. L. Sanders, X. Zhang, F. Y. Yamaner, and O. Oralkan, "An FPGA-based backend system for intravascular photoacoustic and ultrasound imaging," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 66, no. 1, pp. 45–56, Jan. 2019, doi: [10.1109/TUFFC.2018.2881409](https://doi.org/10.1109/TUFFC.2018.2881409).
- [4] S. Rezaei, P. Ghafarian, A. K. Jha, A. Rahmim, S. Sarkar, and M. R. Ay, "Joint compensation of motion and partial volume effects by iterative deconvolution incorporating wavelet-based denoising in oncologic PET/CT imaging," *Phys. Medica*, vol. 68, pp. 52–60, Dec. 2019, doi: [10.1016/j.ejmp.2019.10.031](https://doi.org/10.1016/j.ejmp.2019.10.031).
- [5] C. Huang, Y. Xie, Y. Lan, Y. Hao, F. Chen, Y. Cheng, and Y. Peng, "A new framework for the integrative analytics of intravascular ultrasound and optical coherence tomography images," *IEEE Access*, vol. 6, pp. 36408–36419, 2018, doi: [10.1109/ACCESS.2018.2839694](https://doi.org/10.1109/ACCESS.2018.2839694).
- [6] X. Jin, Q. Jiang, X. Chu, X. Lang, S. Yao, K. Li, and W. Zhou, "Brain medical image fusion using L2-norm-based features and fuzzy-weighted measurements in 2-D Littlewood–Paley EWT domain," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 8, pp. 5900–5913, Aug. 2020, doi: [10.1109/TIM.2019.2962849](https://doi.org/10.1109/TIM.2019.2962849).
- [7] D. Banik, K. Roy, D. Bhattacharjee, M. Nasipuri, and O. Krejcar, "PolypNet: A multimodal fusion network for polyp segmentation," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, 2021, doi: [10.1109/TIM.2020.3015607](https://doi.org/10.1109/TIM.2020.3015607).
- [8] O. S. Faragallah, H. El-Hoseny, W. El-Shafai, W. A. El-Rahman, H. S. El-Sayed, E.-S.-M. El-Rabaie, F. E. A. El-Samie, and G. G. N. Geweid, "A comprehensive survey analysis for present solutions of medical image fusion and future directions," *IEEE Access*, vol. 9, pp. 11358–11371, 2021, doi: [10.1109/ACCESS.2020.3048315](https://doi.org/10.1109/ACCESS.2020.3048315).
- [9] Z. Yan, X. Xu, Y. Wang, T. Li, B. Ma, L. Yang, Y. Lu, and Q. Li, "Application of ultrasonic Doppler technology based on wavelet threshold denoising algorithm in fetal heart rate and central nervous system malformation detection," *World Neurosurgery*, vol. 149, pp. 380–387, May 2021, doi: [10.1016/j.wneu.2020.10.030](https://doi.org/10.1016/j.wneu.2020.10.030).
- [10] B. Ji, Z. Hosseini, L. Wang, L. Zhou, X. Tu, and H. Mao, "Spectral wavelet-feature analysis and classification assisted denoising for enhancing magnetic resonance spectroscopy," *NMR Biomed.*, vol. 34, no. 6, Mar. 2021, Art. no. e4497, doi: [10.1002/nbm.4497](https://doi.org/10.1002/nbm.4497).
- [11] Z. Qu, X. Yan, J. Pan, and P. Chen, "Sparse view CT image reconstruction based on total variation and wavelet frame regularization," *IEEE Access*, vol. 8, pp. 57400–57413, 2020, doi: [10.1109/ACCESS.2020.2982229](https://doi.org/10.1109/ACCESS.2020.2982229).
- [12] C. Rodrigues, Z. M. A. Peixoto, and F. M. F. Ferreira, "Ultrasound image denoising using wavelet thresholding methods in association with the bilateral filter," *IEEE Latin Amer. Trans.*, vol. 17, no. 11, pp. 1800–1807, Nov. 2019, doi: [10.1109/TLA.2019.8986417](https://doi.org/10.1109/TLA.2019.8986417).
- [13] E. Barnhill, L. Hollis, I. Sack, J. Braun, and P. R. Hoskins, "Non-linear multiscale regularisation in MR elastography: Towards fine feature mapping," *Med. Image Anal.*, vol. 35, pp. 133–145, Jan. 2017, doi: [10.1016/j.media.2016.05.012](https://doi.org/10.1016/j.media.2016.05.012).
- [14] J. V. Manjón, P. Coupé, and A. Buades, "MRI noise estimation and denoising using non-local PCA," *Med. Image Anal.*, vol. 22, no. 1, pp. 35–47, May 2015, doi: [10.1016/j.media.2015.01.004](https://doi.org/10.1016/j.media.2015.01.004).
- [15] Z. Xu, M. Gao, G. Z. Papadakis, B. Luna, S. Jain, D. J. Mollura, and U. Bagci, "Joint solution for PET image segmentation, denoising, and partial volume correction," *Med. Image Anal.*, vol. 46, pp. 229–243, May 2018, doi: [10.1016/j.media.2018.03.007](https://doi.org/10.1016/j.media.2018.03.007).
- [16] F. Schirmacher, T. Köhler, J. Endres, T. Lindenberger, L. Husvogt, J. G. Fujimoto, J. Hornegger, A. Dörfler, P. Hoelter, and A. K. Maier, "Temporal and volumetric denoising via quantile sparse image prior," *Med. Image Anal.*, vol. 48, pp. 131–146, Aug. 2018, doi: [10.1016/j.media.2018.06.002](https://doi.org/10.1016/j.media.2018.06.002).
- [17] L. F. R. Lucas, N. M. M. Rodrigues, L. A. da Silva Cruz, and S. M. M. de Faria, "Lossless compression of medical images using 3-D predictors," *IEEE Trans. Med. Imag.*, vol. 36, no. 11, pp. 2250–2260, Nov. 2017, doi: [10.1109/TMI.2017.2714640](https://doi.org/10.1109/TMI.2017.2714640).
- [18] M. Faiza, S. Adabi, B. Daoud, and M. R. N. Avanaki, "High-resolution wavelet-fractal compressed optical coherence tomography images," *Appl. Opt.*, vol. 56, no. 4, p. 1119, Feb. 2017, doi: [10.1364/ao.56.001119](https://doi.org/10.1364/ao.56.001119).
- [19] L. Fang, S. Li, X. Kang, J. A. Izatt, and S. Farsiu, "3-D adaptive sparsity based image compression with applications to optical coherence tomography," *IEEE Trans. Med. Imag.*, vol. 34, no. 6, pp. 1306–1320, Jun. 2015, doi: [10.1109/TMI.2014.2387336](https://doi.org/10.1109/TMI.2014.2387336).
- [20] E. Smistad, A. Ostvik, and A. Pedersen, "High performance neural network inference, streaming, and visualization of medical images using FAST," *IEEE Access*, vol. 7, pp. 136310–136321, 2019, doi: [10.1109/ACCESS.2019.2942441](https://doi.org/10.1109/ACCESS.2019.2942441).
- [21] D. Rossinelli, G. Fourestey, F. Schmidt, B. Busse, and V. Kurtcuoglu, "High-throughput lossy-to-lossless 3D image compression," *IEEE Trans. Med. Imag.*, vol. 40, no. 2, pp. 607–620, Feb. 2021, doi: [10.1109/TMI.2020.3033456](https://doi.org/10.1109/TMI.2020.3033456).
- [22] D. Lanz and A. Kaup, "Graph-based compensated wavelet lifting for scalable lossless coding of dynamic medical data," *IEEE Trans. Image Process.*, vol. 29, pp. 2439–2451, 2020, doi: [10.1109/TIP.2019.2947138](https://doi.org/10.1109/TIP.2019.2947138).
- [23] M. R. Lone, "A high speed and memory efficient algorithm for perceptually-lossless volumetric medical image compression," *J. King Saud Univ.-Comput. Inf. Sci.*, Apr. 2020, doi: [10.1016/j.jksuci.2020.04.014](https://doi.org/10.1016/j.jksuci.2020.04.014).
- [24] D. Dhoubi, A. Naït-Ali, C. Olivier, and M. S. Naceur, "ROI-based compression strategy of 3D MRI brain datasets for wireless communications," *IRBM*, vol. 42, no. 3, pp. 146–153, Jun. 2021, doi: [10.1016/j.irbm.2020.05.001](https://doi.org/10.1016/j.irbm.2020.05.001).
- [25] B. M. G. Rosa and G. Z. Yang, "Bladder volume monitoring using electrical impedance tomography with simultaneous multi-tone tissue stimulation and DFT-based impedance calculation inside an FPGA," *IEEE Trans. Biomed. Circuits Syst.*, vol. 14, no. 4, pp. 775–786, Aug. 2020, doi: [10.1109/TBCAS.2020.3008831](https://doi.org/10.1109/TBCAS.2020.3008831).

- [26] J. Zhang, T. Xu, S. Chen, and X. Wang, "Efficient colorful Fourier ptychographic microscopy reconstruction with wavelet fusion," *IEEE Access*, vol. 6, pp. 31729–31739, 2018, doi: [10.1109/ACCESS.2018.2841854](https://doi.org/10.1109/ACCESS.2018.2841854).
- [27] E. Hahamovich and A. Rosenthal, "Ultrasound detection arrays via coded Hadamard apertures," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 67, no. 10, pp. 2095–2102, Oct. 2020, doi: [10.1109/TUFFC.2020.2993583](https://doi.org/10.1109/TUFFC.2020.2993583).
- [28] L. Cao, H. Li, Y. Zhang, L. Zhang, and L. Xu, "Hierarchical method for cataract grading based on retinal images using improved Haar wavelet," *Inf. Fusion*, vol. 53, pp. 196–208, Jan. 2020, doi: [10.1016/j.inffus.2019.06.022](https://doi.org/10.1016/j.inffus.2019.06.022).
- [29] S. Wang, J. Lv, Z. He, D. Liang, Y. Chen, M. Zhang, and Q. Liu, "Denoising auto-encoding priors in undecimated wavelet domain for MR image reconstruction," *Neurocomputing*, vol. 437, pp. 325–338, May 2021, doi: [10.1016/j.neucom.2020.09.086](https://doi.org/10.1016/j.neucom.2020.09.086).
- [30] D. D. Sundararajan, *Digital Signal Processing*. Cham, Switzerland: Springer, 2021.
- [31] S. V. Renuka and D. R. Edla, "Adaptive shrinkage on dual-tree complex wavelet transform for denoising real-time MR images," *Biocybern. Biomed. Eng.*, vol. 39, no. 1, pp. 133–147, Jan. 2019, doi: [10.1016/j.bbe.2018.11.003](https://doi.org/10.1016/j.bbe.2018.11.003).
- [32] R. Boujelbene, L. Boubchir, and Y. B. Jemaa, "Enhanced embedded zerotree wavelet algorithm for lossy image coding," *IET Image Process.*, vol. 13, no. 8, pp. 1364–1374, Jun. 2019, doi: [10.1049/iet-ipr.2018.6052](https://doi.org/10.1049/iet-ipr.2018.6052).
- [33] M. Mardani, E. Gong, J. Y. Cheng, S. S. Vasanaawala, G. Zaharchuk, L. King, and J. M. Pauly, "Deep generative adversarial neural Networks for compressive sensing MRI," *IEEE Trans. Med. Imaging*, vol. 38, no. 1, pp. 167–179, Jan. 2019, doi: [10.1109/TMI.2018.2858752](https://doi.org/10.1109/TMI.2018.2858752).
- [34] M. Ravi, A. Sewa, T. G. Shashidhar, and S. S. S. Sanagapati, "FPGA as a hardware accelerator for computation intensive maximum likelihood expectation maximization medical image reconstruction algorithm," *IEEE Access*, vol. 7, pp. 111727–111735, 2019, doi: [10.1109/ACCESS.2019.2932647](https://doi.org/10.1109/ACCESS.2019.2932647).
- [35] J. Janjic, M. Tan, V. Daeichin, E. Noothout, C. Chen, Z. Chen, and Z.-Y. Chang, "A 2-D ultrasound transducer with front-end ASIC and low cable count for 3-D forward-looking intravascular imaging: Performance and characterization," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 65, no. 10, pp. 1832–1844, Oct. 2018, doi: [10.1109/TUFFC.2018.2859824](https://doi.org/10.1109/TUFFC.2018.2859824).
- [36] X. Zhai, A. Amira, F. Bensaali, A. Al-Shibani, A. Al-Nassr, A. El-Sayed, M. Eslami, S. P. Dakua, and J. Abinshed, "Zynq SoC based acceleration of the lattice Boltzmann method," *Concurrency Comput., Pract. Exp.*, vol. 31, no. 17, p. e5184, Sep. 2019, doi: [10.1002/cpe.5184](https://doi.org/10.1002/cpe.5184).
- [37] A. Mehrnia and A. N. Willson, "A lower bound for the hardware complexity of FIR filters," *IEEE Circuits Syst. Mag.*, vol. 18, no. 1, pp. 10–28, Jan. 2018, doi: [10.1109/MCAS.2017.2785422](https://doi.org/10.1109/MCAS.2017.2785422).
- [38] H. Y. Alzaq and B. B. Ustundag, "An optimized two-level discrete wavelet implementation using residue number system," *EURASIP J. Adv. Signal Process.*, vol. 2018, no. 1, p. 41, Dec. 2018, doi: [10.1186/s13634-018-0559-3](https://doi.org/10.1186/s13634-018-0559-3).
- [39] A. S. Molahosseini, L. S. de Seabra, and C. H. Chang, *Embedded Systems Design With Special Arithmetic and Number Systems*. Cham, Switzerland: Springer, 2017.
- [40] P. V. A. Mohan, *Residue Number Systems: Theory and Applications*. Cham, Switzerland: Springer, 2016.
- [41] R. Krishnaswamy, "VLSI based orthogonal diagonal cross hair search (ODCHS) algorithm implementation for efficient image compression," *Microprocessors Microsystems*, vol. 76, Jul. 2020, Art. no. 103114, doi: [10.1016/j.micpro.2020.103114](https://doi.org/10.1016/j.micpro.2020.103114).
- [42] A. Ibrahim, S. Zhang, F. Angiolini, M. Arditì, S. Kimura, S. Goto, J.-P. Thiran, and G. De Micheli, "Towards ultrasound everywhere: A portable 3D digital back-end capable of zone and compound imaging," *IEEE Trans. Biomed. Circuits Syst.*, vol. 12, no. 5, pp. 968–981, Oct. 2018, doi: [10.1109/TBCAS.2018.2828382](https://doi.org/10.1109/TBCAS.2018.2828382).
- [43] G. D. Licciardo, C. Cappetta, and L. Di Benedetto, "Design of a Gabor filter HW accelerator for applications in medical imaging," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 8, no. 7, pp. 1187–1194, Jul. 2018, doi: [10.1109/TCPMT.2018.2818947](https://doi.org/10.1109/TCPMT.2018.2818947).
- [44] S. G. Sreejeesh, R. Sakthivel, and J. U. Kidav, "Superior implementation of accelerated QR decomposition for ultrasound imaging," *IEEE Access*, vol. 8, pp. 156244–156260, 2020, doi: [10.1109/ACCESS.2020.3017558](https://doi.org/10.1109/ACCESS.2020.3017558).
- [45] S.-N. Tang and F.-C. Jan, "Energy-efficient and calibration-aware Fourier-domain OCT imaging processor," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 27, no. 6, pp. 1390–1403, Jun. 2019, doi: [10.1109/TVLSI.2019.2894751](https://doi.org/10.1109/TVLSI.2019.2894751).
- [46] E. Kang, Q. Ding, M. Shabanmotlagh, P. Kruijzinga, Z.-Y. Chang, E. Noothout, H. J. Vos, J. G. Bosch, M. D. Verweij, N. de Jong, and M. A. P. Pertijs, "A reconfigurable ultrasound transceiver ASIC with 24×40 elements for 3-D carotid artery imaging," *IEEE J. Solid-State Circuits*, vol. 53, no. 7, pp. 2065–2075, Jul. 2018, doi: [10.1109/JSSC.2018.2820156](https://doi.org/10.1109/JSSC.2018.2820156).
- [47] V. Nadig, B. Weissler, H. Radermacher, V. Schulz, and D. Schug, "Investigation of the power consumption of the PETsys TOFPET2 ASIC," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 4, no. 3, pp. 378–388, May 2020, doi: [10.1109/TRPMS.2019.2955032](https://doi.org/10.1109/TRPMS.2019.2955032).
- [48] M. Occhipinti, M. Carminati, P. Busca, A. D. Butt, G. L. Montagnani, P. Trigilio, C. Piemonte, A. Ferri, A. Gola, T. Bukki, M. Czeller, Z. Nyitrai, Z. Papp, K. Nagy, and C. Fiorini, "Characterization of the detection module of the INSERT SPECT/MRI clinical system," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 2, no. 6, pp. 554–563, Nov. 2018, doi: [10.1109/TRPMS.2018.2864792](https://doi.org/10.1109/TRPMS.2018.2864792).
- [49] P. Trigilio, P. Busca, R. Quaglia, M. Occhipinti, and C. Fiorini, "A SiPM-readout ASIC for SPECT applications," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 2, no. 5, pp. 404–410, Feb. 2018, doi: [10.1109/TRPMS.2018.2856201](https://doi.org/10.1109/TRPMS.2018.2856201).
- [50] D. Baudin, S. Dubos, O. Gevin, O. Limousin, D. Maier, A. Michalowska, D. Renaud, P. Serrano, T. Takahashi, and S. Watanabe, "D2R1: A 2-D X-ray detector for CdTe-based fine pitch and high-energy resolution imaging spectroscopy," *IEEE Trans. Nucl. Sci.*, vol. 65, no. 7, pp. 1408–1415, Jul. 2018, doi: [10.1109/TNS.2018.2845461](https://doi.org/10.1109/TNS.2018.2845461).
- [51] M. D. Adams and F. Kossentini, "Reversible integer-to-integer wavelet transforms for image compression: Performance evaluation and analysis," *IEEE Trans. Image Process.*, vol. 9, no. 6, pp. 1010–1024, Jun. 2000, doi: [10.1109/83.846244](https://doi.org/10.1109/83.846244).
- [52] D. M. B. Larrotta, D. M. M. Enciso, and A. E. G. Barrera, "Compression of biomedical signals on FPGA by DWT and run-length," in *Proc. IEEE ANDESCON*, Sep. 2010, pp. 1–5, doi: [10.1109/ANDESCON.2010.5633621](https://doi.org/10.1109/ANDESCON.2010.5633621).
- [53] M. Chehaitly, M. Tabaa, F. Monteiro, and A. Dandache, "A fast and configurable architecture for discrete wavelet packet transform," in *Proc. Conf. Des. Circuits Integr. Syst. (DCIS)*, Nov. 2015, pp. 1–6, doi: [10.1109/DCIS.2015.7388599](https://doi.org/10.1109/DCIS.2015.7388599).
- [54] M. Chehaitly, M. Tabaa, F. Monteiro, and A. Dandache, "A ultra high speed and configurable inverse discrete wavelet packet transform architecture," in *Proc. 29th Int. Conf. Microelectron. (ICM)*, Dec. 2017, pp. 1–4, doi: [10.1109/ICM.2017.8268863](https://doi.org/10.1109/ICM.2017.8268863).
- [55] N. I. Chervyakov, P. A. Lyakhov, and N. N. Nagornov, "Quantization noise of multilevel discrete wavelet transform filters in image processing," *Optoelectron., Instrum. Data Process.*, vol. 54, no. 6, pp. 608–616, Nov. 2018, doi: [10.3103/S8756699018060092](https://doi.org/10.3103/S8756699018060092).
- [56] B. Parhami, *Computer Arithmetic: Algorithms and Hardware Designs*. Oxford, U.K.: Oxford Univ. Press, 2010.
- [57] A. R. Omondi and B. Premkumar, *Residue Number Systems: Theory and Implementation*. London, U.K.: Imperial College Press, 2007.
- [58] A. Safari, C. V. Niras, and Y. Kong, "Power-performance enhancement of two-dimensional RNS-based DWT image processor using static voltage scaling," *Integration*, vol. 53, pp. 145–156, Mar. 2016, doi: [10.1016/j.vlsi.2015.12.006](https://doi.org/10.1016/j.vlsi.2015.12.006).
- [59] N. I. Chervyakov, P. A. Lyakhov, A. S. Ionisyan, and M. V. Valueva, "High-speed smoothing filter in the residue number system," in *Proc. 3rd Int. Conf. Digit. Inf. Process., Data Mining, Wireless Commun. (DIPDMWC)*, Jul. 2016, pp. 121–126, doi: [10.1109/DIPDMWC.2016.7529375](https://doi.org/10.1109/DIPDMWC.2016.7529375).
- [60] P. Patronik and S. J. Piestrak, "Hardware/software approach to designing low-power RNS-enhanced arithmetic units," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 64, no. 5, pp. 1031–1039, May 2017, doi: [10.1109/TCASI.2017.2669108](https://doi.org/10.1109/TCASI.2017.2669108).
- [61] H. Alzaq and B. B. Üstündag, "Multiplier-less 1-level discrete wavelet transform implementations on ZC706 development kit," in *Proc. 10th Int. Conf. Elect. Electron. Eng.*, 2017, pp. 1122–1126.
- [62] N. I. Chervyakov, P. A. Lyakhov, N. N. Nagornov, D. I. Kaplun, A. S. Voznesenskiy, and D. V. Bogayevskiy, "Implementation of smoothing image filtering in the residue number system," in *Proc. 8th Medit. Conf. Embedded Comput. (MECO)*, Jun. 2019, pp. 1–4, doi: [10.1109/MECO.2019.8760190](https://doi.org/10.1109/MECO.2019.8760190).

- [63] P. Lyakhov, M. Valueva, G. Valuev, and N. Nagornov, "High-performance digital filtering on truncated multiply-accumulate units in the residue number system," *IEEE Access*, vol. 8, pp. 209181–209190, 2020, doi: [10.1109/ACCESS.2020.3038496](https://doi.org/10.1109/ACCESS.2020.3038496).
- [64] N. I. Chervyakov, P. A. Lyakhov, N. N. Nagornov, M. V. Valueva, and D. I. Kaplun, "High-performance hardware 3D medical imaging using wavelets in the residue number system," in *Proc. 9th Medit. Conf. Embedded Comput. (MECO)*, Jun. 2020, pp. 1–4, doi: [10.1109/MECO49872.2020.9134123](https://doi.org/10.1109/MECO49872.2020.9134123).
- [65] B. Mohindroo, A. Paliwal, and K. Suneja, "FPGA based faster implementation of MAC unit in residual number system," in *Proc. Int. Conf. for Emerg. Technol. (INCET)*, Jun. 2020, pp. 1–4, doi: [10.1109/INCET49848.2020.9154105](https://doi.org/10.1109/INCET49848.2020.9154105).
- [66] M. V. Valueva, N. N. Nagornov, P. A. Lyakhov, G. V. Valuev, and N. I. Chervyakov, "Application of the residue number system to reduce hardware costs of the convolutional neural network implementation," *Math. Comput. Simul.*, vol. 177, pp. 232–243, Nov. 2020, doi: [10.1016/j.matcom.2020.04.031](https://doi.org/10.1016/j.matcom.2020.04.031).
- [67] N. I. Chervyakov, P. A. Lyakhov, M. A. Deryabin, N. N. Nagornov, M. V. Valueva, and G. V. Valuev, "Residue number system-based solution for reducing the hardware cost of a convolutional neural network," *Neurocomputing*, vol. 407, pp. 439–453, Sep. 2020, doi: [10.1016/j.neucom.2020.04.018](https://doi.org/10.1016/j.neucom.2020.04.018).
- [68] D. G. Bailey, *Design for Embedded Image Processing on FPGAs*. Singapore: Wiley, 2011.
- [69] K. R. Rao and P. C. Yip, *The Transform and Data Compression Handbook*. Boca Raton, FL, USA: CRC Press, 2001.
- [70] S. Lalithakumari, R. Pandian, J. Rani, D. Vinothkumar, A. Sneha, V. Amalarani, and B. Joe, "Selection of optimum compression algorithms based on the characterization on feasibility for medical image," *Biomed. Res.*, vol. 28, no. 13, pp. 5633–5637, 2017.
- [71] H. T. Vergos and G. Dimitrakopoulos, "On modulo  $2^{n+1}$  adder design," *IEEE Trans. Comput.*, vol. 61, no. 2, pp. 173–186, Feb. 2012, doi: [10.1109/TC.2010.261](https://doi.org/10.1109/TC.2010.261).
- [72] D. Zivaljevic, N. Stamenkovic, and V. Stojanovic, "Digital filter implementation based on the RNS with diminished-1 encoded channel," in *Proc. 35th Int. Conf. Telecommun. Signal Process. (TSP)*, Jul. 2012, pp. 662–666, doi: [10.1109/TSP.2012.6256380](https://doi.org/10.1109/TSP.2012.6256380).
- [73] R. de Matos, R. Paludo, N. Chervyakov, P. A. Lyakhov, and H. Pettenghi, "Efficient implementation of modular multiplication by constants applied to RNS reverse converters," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2017, pp. 1–4, doi: [10.1109/ISCAS.2017.8050779](https://doi.org/10.1109/ISCAS.2017.8050779).
- [74] P. M. Kogge and H. S. Stone, "A parallel algorithm for the efficient solution of a general class of recurrence equations," *IEEE Trans. Comput.*, vol. COM-22, no. 8, pp. 786–793, Aug. 1973, doi: [10.1109/TC.1973.5009159](https://doi.org/10.1109/TC.1973.5009159).
- [75] 1076–2019—*IEEE Standard for VHDL Language Reference Manual | IEEE Standard | IEEE Xplore*. Accessed: Nov. 23, 2021. [Online]. Available: <https://ieeexplore.ieee.org/document/8938196>
- [76] N. I. Chervyakov, P. A. Lyakhov, D. I. Kalita, and K. S. Shulzhenko, "Effect of RNS dynamic range on grayscale images filtering," in *Proc. XV Int. Symp. Problems Redundancy Inf. Control Syst.*, Sep. 2016, pp. 33–37, doi: [10.1109/RED.2016.7779323](https://doi.org/10.1109/RED.2016.7779323).



**NIKOLAY N. NAGORNOV** received the bachelor's and master's degrees in applied mathematics and computer science from North-Caucasus Federal University (NCFU), in 2014 and 2016, respectively, and the Ph.D. degree in computer science in 2021. He works with the Department of Mathematical Modeling, NCFU. His research interests include modular arithmetic, digital image processing, and medical image processing.



**PAVEL A. LYAKHOV** received the Graduate and Ph.D. degrees in mathematics from Stavropol State University, in 2009 and 2012, respectively. He has been working with North-Caucasus Federal University, since 2012, where he is currently the Head of the Department of Mathematical Modeling. He also works with the North-Caucasus Center for Mathematical Research as the Head of the Department of Modular Computations and Artificial Intelligence. He has authored over 50 scientific publications. His research interests include high-performance computing, residue number systems, digital signal and image processing, and machine learning.



**MARIA V. VALUEVA** (Graduate Student Member, IEEE) received the bachelor's and master's degrees in applied mathematics and computer science from North-Caucasus Federal University (NCFU), in 2016 and 2018, respectively, where she is currently pursuing the Ph.D. degree. She works with the North-Caucasus Center for Mathematical Research, NCFU, as a Junior Researcher. Her research interests include high-performance computing, modular arithmetic, artificial neural networks, and digital image processing.



**MAXIM V. BERGERMAN** (Graduate Student Member, IEEE) received the bachelor's and master's degrees in applied mathematics and computer science from North-Caucasus Federal University (NCFU), in 2019 and 2021, respectively, where he is currently pursuing the Ph.D. degree. He works with the North-Caucasus Center for Mathematical Research, Department of Modular Computations and Artificial Intelligence, NCFU, as a Junior Researcher. His research interests include high-performance computing, modular arithmetic, and digital signal processing.

...