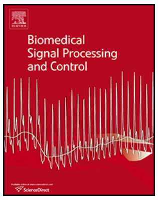
Biomedical Signal Processing and Control 8 (2013) 682–687

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc



Short communication

Context modeling based lossless compression of radio-frequency data for software-based ultrasound beamforming

Chai-eun Lim[a], Gi-duck Kim[b], Changhan Yoon[a], Jung-jun Kim[a], Kang-sik Kim[c], Tai-kyong Song[a,*], Yangmo Yoo[a,*]

[a] Department of Electronic Engineering, Sogang University, Seoul, Republic of Korea
[b] Sogang Institute of Advanced Technology, Sogang University, Seoul, Republic of Korea
[c] Samsung Electronics, Suwon, Republic of Korea

A R T I C L E   I N F O

Article history:
Received 2 April 2012
Received in revised form 10 June 2013
Accepted 14 June 2013
Available online 30 July 2013

Keywords:
Ultrasound beamforming
Lossless compression
Context modeling
Entropy coding

A B S T R A C T

A new lossless compression method using context modeling for ultrasound radio-frequency (RF) data is presented. In the proposed compression method, the combination of context modeling and entropy coding is used for effectively lowering the data transfer rates for modern software-based medical ultrasound imaging systems. From the phantom and *in vivo* data experiments, the proposed lossless compression method provides the average compression ratio of 0.45 compared to the Burg and JPEG-LS methods (0.52 and 0.55, respectively). This result indicates that the proposed compression method is capable of transferring 64-channel 40-MHz ultrasound RF data with a 16-lane PCI-Express 2.0 bus for software beamforming in real time.

1. Introduction

1.1. Medical imaging and technology trend

In the modern medicine, medical imaging equipment has been increasingly playing an important role in screening, diagnosis, treatment and follow-up of lesions. Especially, the image quality of the most widely used diagnostic ultrasound imaging systems has been considerably improved by taking advantages of advance in electronic technology mainly driven by semiconductor industry.

The technological advance in single core microprocessors also drives rapid performance improvement and substantial cost reduction in high-performance computing applications. This effort, however, has been switching to multi-core and multi-thread models since single-core processors have reached to physical limitation to further increase the clock speed due to power consumption and heat dissipation issues.

Unlike a central processing unit (CPU), a graphics processing unit (GPU) [1], one of the parallel processors, allots more die to an arithmetic-logic unit (ALU) than cache, leading to substantial enhancement in computational power. There have been many attempts to apply the high-performance GPU to the general computing. This trend is recently accelerated with the emergence of the new programming model, such as Compute Unified Device Architecture (CUDA) [2] and Open Computing Language (OpenCL) [3]. Similarly, there are attempts to handle ultrasound signal and image processing with software, which has traditionally been handled by hardware.

In spite of availability of super-computing power on multi-processors, it has been debated whether ultrasound receive beamforming can be executed in software because of huge amount data and computation to handle. Although computational hurdles could be greatly diminished as high-performance multi-core processors develop, transfer bandwidth for ultrasound radio-frequency (RF) data from analog-to-digital converters (ADCs) to a PC still remains bottleneck.

1.2. Ultrasound RF data bandwidth

If an ultrasound frame consists of 128 vectors and the view depth is 15 cm, 10-bit, 40-MHz, 64-channel RF data for the frame rate of 30 Hz are given by

$$\begin{aligned}&128\,(\text{vectors}) \times 64\,(\text{channels}) \times 2\,(\text{bytes})\\ &\quad\times 8192\,(\text{samples}) \times 30\,(\text{Hz}) = 3.84\,\text{GB}\end{aligned} \tag{1}$$

As indicated in Eq. (1), the commonly used PCI-Express 2.0 bandwidth is not sufficient for the transmission of these RF data in real time; thus effective compression for ultrasound RF data is necessary

---

* Corresponding authors at: Sogang University, 1 Shinsu-dong, Mapo-gu, Seoul 121-742, Republic of Korea. Tel.: +82 2 705 4731; fax: +82 2 707 3008.
E-mail addresses: tksong@sogang.ac.kr (T.-k. Song), ymyoo@sogang.ac.kr (Y. Yoo).

1746-8094/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bspc.2013.06.010

for software beamforming running on a GPU. As the computational capability is being improved more than it is for the transfer bandwidth, it is appropriate to utilize the remaining computational power to encode, transmit and decode the RF data.

1.3. Lossless compression of ultrasound RF data

Image compression has attracted more interests from the medical service sector. It can be achieved by lossy or lossless compression. Lossless compression is generally preferred in medical imaging to ensure a high resolution image and legal safeguard. Lossy compression requires extensive evaluation on how much loss is acceptable clinically, but lossless compression does not since it preserves the original data.

In general, data can be compressed by removing redundancy (spatial, temporal and statistical) between adjacent symbols. Spatial and temporal redundancy can be removed by prediction. Similarly, statistical redundancy can be removed by context modeling of the error during entropy encoding. Moreover, redundancy hidden in the original data may vary depending on the data; thus, compression needs to accommodate the attributes of the data. Similarly, ultrasound RF data have their own characteristics, which need to be taken into account in order to enhance compression rate. For lossless compression, context-based adaptive lossless image coding (CALIC) and JPEG-LS methods were previously proposed [4]. Since CALIC and JPEG-LS are both based on general 2D images, the compression efficiency for ultrasound RF data would be limited. For ultrasound RF data compression, JPEG/JPEG2000 and cross-correlation compression methods were evaluated [5,6]. The JPEG/JPEG2000 methods can provide high compression ratio for ultrasound RF data. However, these lossy compression methods yielded an average error of lower than 5 dB so that they are not appropriate for ultrasound RF data compression. By computing the cross-correlation between current and previous frame's RF data to find optimal shift, the compression based on the residual signal can be performed [6]. This method can provide high compression ratio with the stationary images, but it requires a huge amount of frame memories to store RF data (e.g., >3.84 GB). In addition, this method needs computationally expensive cross-correlation processing so that it is still challenging to be implemented in real time in the modern ultrasound machines. In this paper, a new lossless data compression algorithm for real-time transfer of ultrasound RF data with the PCI-Express 2.0 bandwidth is presented and its performance is compared with two lossless compression techniques (i.e., Burg and JPEG-LS).

2. Materials and methods

2.1. Flowchart for the proposed context-based lossless compression method for ultrasound RF data

The flowchart for the proposed context-based lossless compression (CBLC) method for ultrasound RF data, which consists of context-based prediction and binary arithmetic coding with bit-plane decomposition, is illustrated in Fig. 1.

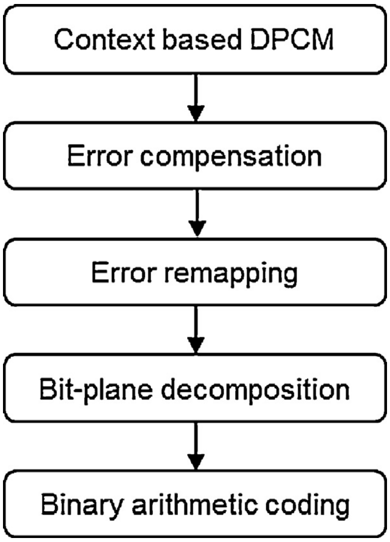


Fig. 1. Flowchart for the proposed context-based lossless compression method.

2.2. Context-based prediction

In general, prediction can be performed based on both horizontal and vertical gradients in CALIC [4], simple conditional expression in JPEG-LS or their combination [7]. In addition to non-linear methodology of going through conditional expression based on gradients, linear prediction using auto-regression (AR) can be applied [8,9]. Linear prediction is a mathematical operation where further values of a discrete-time signal are estimated as a linear function of previous samples as follows:

$$\hat{x}(n) = \sum_{i=1}^{p} a_i x(n-i) \qquad (2)$$

where p is the order of the AR model and a_i is the AR coefficient that can be found by using previous data based on the Levinson–Dubin or Burg algorithm [12]. Although such AR prediction is fairly simple if the coefficient is provided, it requires additional computations to find the coefficients in real time. This is why the Burg method is not suitable for the real-time implementation. In the proposed CBLC method, the prediction based on a simple differential pulse code modulation (DPCM) ($p = 1$ in Eq. (2)) is adopted for AR modeling as like

$$\hat{x}(n) = x(n-1) \qquad (3)$$

The prediction alone cannot exploit the spatial characteristics of the symbols. In CBLC, context refers to the samples prior to the current symbol and the performance of a predictor can be substantially improved by context modeling. For example, with10-bit ADCs, ultrasound RF data vary from 0 to 1023 when they are converted to an unsigned value. If all 1024 values are used for context modeling, there are too many contexts. In other words, even if only three contexts are used to predict the current value, the total size of the context table becomes as much as 1024 × 1024 × 1024. In addition to this tremendous memory requirement, the large context can lead to the sparse context. To overcome these problems, 1024 values can be divided into the N numbers of blocks to decrease the quantization error of context [15]. In this paper, N is chosen as eight and only the top three bits from the 10-bit RF data are used. Thus, if three contexts are used, making the whole context table becomes nine bits, i.e., total 512. Table 1 shows the example of context construction for data A, B, C, D, and E and their context mapping where a, b, c, d, and e are the top 3 bits of A, B, C, D, and E.

For the illustrative example, a Gaussian pulse and its context table are shown in Fig. 2 and Table 2, respectively. The Gaussian pulse was generated with a sampling and center frequencies of 40 MHz and 3 MHz, respectively, and a fractional bandwidth of the pulse was 0.7. The pulse was quantized to 10-bit unsigned values. The characteristic of the generated pulse with these parameters is similar to that used in a typical ultrasound imaging system. In this example, the context table was constructed with only half of pulse since the shape of pulse is symmetry and three contexts were also utilized; thus, context table consisted with nine bits.