PCM-based Architecture for Compressed Sensing on Skin Ulcers Images and Automatic Classification with CNN Neural Network

Original
PCM-based Architecture for Compressed Sensing on Skin Ulcers Images and Automatic Classification with CNN Neural Network / Cavazzana, R., Secco, J., Pareschi, F., Corinto, F.. - STAMPA. - (2024). (2024 IEEE International Flexible Electronics Technology Conference (IFETC2024) Bologna (Ita) 15-18 September 2024) [10.1109/IFETC61155.2024.10771843].

Availability:
This version is available at: 11583/2995723 since: 2024-12-20T10:49:51Z

Publisher:
IEEE

Published
DOI:10.1109/IFETC61155.2024.10771843

(Article begins on next page)

This is the author's version of the article that has been presented at IEEE IFETC2024
The editorial version of the paper is available at http://dx.doi.org/10.1109/IFETC61155.2024.10771843

PCM-based Architecture for Compressed Sensing on Skin Ulcers Images and Automatic Classification with CNN Neural Network

Rosanna Cavazzana
Department of Electronics and Telecommunications
Politecnico di Torino
Turin, Italy
rosanna.cavazzana@polito.it

Jacopo Secco
Department of Electronics and Telecommunications
Politecnico di Torino
Turin, Italy
jacopo.secco@polito.it

Fabio Pareschi
Department of Electronics and Telecommunications
Politecnico di Torino
Turin, Italy
fabio.preschi@polito.it

Fernando Corinto
Department of Electronics and Telecommunications
Politecnico di Torino
Turin, Italy
fernando.coirinto@polito.it

***Abstract*—Phase Change Memory (PCM) represents a technology that exploits the reversible phase transition of a chalcogenide material to create nanoscale memory components, which can be used for the development of brain-inspired computing approaches. These PCM devices have been examined both as non-volatile storage-class memory and as computing elements for in-memory and neuromorphic computing applications. It is also known that PCM exhibits several characteristics of a memristive device. In this study, we consider a PCM array as an encoder within a system for applying Compressed Sensing (CS) to images of skin ulcers. We then use a decoding strategy that compensates for the non-linearity of PCM devices through an iterative optimization approach. The quality of image reconstruction was evaluated by classifying the images using a Convolutional Neural Network (CNN) according to Wound Bed Preparation (WBP) severity scale, which is used in clinical practice for the assessment of skin lesions. The effectiveness of the image compression and reconstruction was demonstrated by comparing the automatic classification performance on before and after CS images.**

***Index Terms*—PCM, Memristor, Compressed Sensing, Telemedicine, Wound Care, CNN**

I. Introduction

A skin ulcer is a pathological condition that manifests as a chronic wound that has not healed or cannot completely restore its anatomical and functional integrity. This condition is most found in individuals over 65 years of age and affects about 1-2% of the global population. Rapid and precise assessment of the lesion is crucial for defining the correct treatment plan and promoting healing. For this reason, the Wound Viewer (WV) was developed by Omnidermal Biomedics s.r.l., a telemedical device capable of capturing an image of the wound through its camera and automatically classifying it according to the Wound Bed Preparation (WBP) scale using an integrated Artificial Intelligence algorithm. WBP scale is a common skin wound assessment methodology used in clinic to assess its pathological state, i.e. to understand whether the wound is healing, or the tissue is undergoing necrosis [1].

The idea of applying a compression algorithm to the images acquired through the WV arises from the need to reduce the costs of transmitting and storing such images, thereby making data sharing more efficient and reducing storage costs on cloud platforms. In this context, we introduce the Compressed Sensing (CS) theory which has been widely used in the literature for acquiring and reconstructing signals from a number of scalars, called measurements, that is smaller than in a traditional system based on the Nyquist theorem. The encoding step, in which these measurements are obtained, can be represented as a matrix-vector product that can be implemented using a Phase Change Memory (PCM) array.

PCM devices encode information through the phase configuration characterized by a layer of material placed between two metal electrodes. This type of material shows high conductivity in the crystalline state (SET) and much lower conductivity in the amorphous state (RESET). In our work, we considered the Encoder model shown in Fig. 1, which was used and validated in a recent study by C. Antolini et al. for signal acquisition and reconstruction [2].

The authors demonstrated the possibility of using analog elements to perform matrix operations by exploiting Ohm's and Kirchhoff's laws, applying PCM technology for Analog In-Memory Computing (AIMC). We have therefore used the same model, considering that a digital image is nothing more than a signal represented in a two-dimensional space rather than a one-dimensional one. According to this model, given a cell with conductance g and applying an input voltage v to the cell, an output current is read as a single multiplication $i = gv$. By combining the outputs of multiple cells, each characterized by a conductance $g_{k,j}$ and its own applied voltage v_j, a sum

This is the author's version of the article that has been presented at IEEE IFETC2024
The editorial version of the paper is available at http://dx.doi.org/10.1109/IFETC61155.2024.10771843

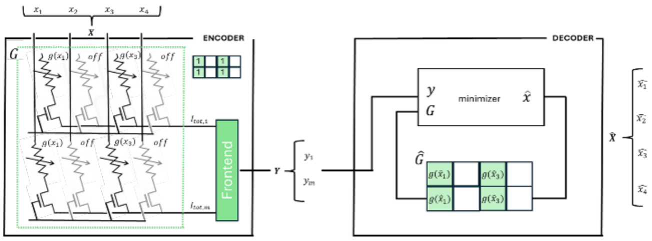

Fig. 1. Architecture of a PCM-based system for CS and decoding applications. Adapted from [2].

of products $I_{TOT,k}$ is obtained, where $j = 1, 2, 3, \ldots, n$ and $k = 1, 2, 3, \ldots, m$.. The entire memory can thus be assimilated to a conductance matrix G of $m \times n$ dimensions. By applying a voltage vector to each row, a matrix-vector multiplication is achieved [3].

For the reconstruction, we adopted an iterative technique that has been shown to be capable of reconstructing a signal even when the measurements are produced by highly variable conductances. For the current application, at each iteration, the input signal is estimated, yielding an improved representation of the signal-dependent conductance matrix. Then this updated matrix is used in the subsequent iteration until convergence is reached.

II. PCM Technology

PCM devices are characterized by a nonlinear dynamic behavior resulting from a complex interaction between thermal, electrical and structural dynamics [4], [5]. These dynamics have been described as a memristive system, introduced in 1976 by S. Kang and L. O. Chua [6], in a previous work by F. Marrone et al. [7], showing the characteristic current-voltage loop of PCM devices when subjected to a periodic bipolar input. This behavior clearly classifies them within the category of memristive systems.

In the previous work by Antolini et al. [2], three different models, labelled 1, 2 and 3, were obtained by programming 5120 cells with different SET pulse intensities and characterizing them in the i/v domain. The conductance model type 2 proved to be the most performant and was therefore used for the CS performed in the current work.

Each instance of the vectorized image is encoded using a matrix based on a PCM, whose cells in the SET state are described by the selected model.

III. Compressed Sensing

CS [8], [9] allows for signal compression through a low-complexity linear encoding phase, which can be considered a simple matrix-vector multiplication. The complexity is shifted to the decoding phase, where an optimization problem must be solved to reconstruct the signal.C S can be applied to an n-dimensional signal $x \in \mathbb{R}^n$ only if it is sparse when represented in any vector space. Given $x = \Psi s$, where $\Psi \in \mathbb{R}^{nxn}$ represents an orthogonal *sparsity basis*, i.e. it satisfies the condition $\Psi\Psi_T = I_n$, where I_n is the $n \times n$ identity matrix, and $\Psi_T\Psi = I, x$ is defined as sparse if only a small fraction of the coefficients in s are significantly non zero. If this condition is met, the CS process can be expressed as $y = Ax$, where $y \in \mathbb{R}^m$ represents the measurements vector and $A \in \mathbb{R}^{mxn}$ is called the measurement matrix. Therefore, to represent x, m measurements $y_1, y_2 \ldots, y_m$ with $m \ll n$ are required, each calculated as $y_k = \sum_{(j=1)}^{n} a_{kj}x_j$.

The reconstruction of the original signal $\hat{x}$ can be found by seeking the sparsest among all signals that generate the same y, through solving an l_0 minimization problem that can be expressed as:

$$\hat{x} = \Psi * argmin_s |s|_1 \quad \text{s.t.} \quad A\Psi s = y \tag{1}$$

Where $|\cdot|_0$ is the l_1 pseudonorm, counting the number of non-zero elements of its argument. The decoding phase therefore requires the exact knowledge of the matrix A, which is impossible due to the voltage dependence of the conductance of PCM cells, resulting in $y = A(x)x$. Even if the exact dependency is known, the decoder cannot derive the matrix $A(x)$ since x is unknown. Previous studies have already conducted preliminary investigations regarding the decoding of a signal after non-linear sensing [10], [11]. However, in the current work, we have adopted the method proposed by C. Paolino et al., which has shown promising results [2].

IV. Methods

Given the three conductance models we utilized the model in state 2. Each instance of the input signal is encoded through the model describing the SET state of the PCM-based array. From the corresponding measurement vector y the input signal is reconstructed, knowing the conductance model and iteratively using it to estimate the actual conductances of the encoder.

Given the ideal matrix $A \in \{0, 1\}^{mxn}$, the measurement vector y, a reference value x_{ref} and the model describing the voltage dependency of the conductance elements $g(x) : \mathbb{R} \rightarrow \mathbb{R}$, the initial conductance matrix is estimated as $\hat{G}_{|0} = Ag(x_{ref})$. This matrix is used to compute $\hat{x}_{|1}$, which is the first estimate of the input signal. The conductance matrix is then updated by recalculating all its elements using the first estimate of the input signal and used to compute a new estimate of the signal $\hat{x}_{|2}$. The algorithm is applied until convergence of the estimated signals $\hat{x}_{|p}$, which inevitably leads to the convergence of the conductance matrices $\hat{G}_{|p}$.

The compression and reconstruction technique were applied to a total of 446 images of skin ulcers, using the Discrete Cosine Transform (DCT). To represent the images in a one-dimensional space, we divided the images into blocks, then vectorized and processed each block individually using a Block CS approach. The technique was implemented with $n = 144$ and a sparsity level of 19 non-zero coefficients. The measurement matrix for each block is binary and with 72×144 dimensions, with ideal zeros and ones implemented by conductance model 2. Then, the blocks were reconstructed and assembled into an image.

This is the author's version of the article that has been presented at IEEE IFETC2024
The editorial version of the paper is available at http://dx.doi.org/10.1109/IFETC61155.2024.10771843

With the reconstructed images after BCS, we performed automatic detection and classification of the acquired lesions. By comparing the classification performance obtained on the original images and on those estimated after CS, it is possible to numerically evaluate the quality of the reconstruction and the relevant information content for classification that is retained. For the classification phase, we utilized a CNN trained on a balanced Training Set of 200 images. We used a Validation Set of 119 images the network parameters tuning and finally a Test Set of 127 images for validating the results. For the calculation of wound detection and classification performance, we used Precision and Recall metrics for each object class to be recognized:

$$\text{Precision}_i = \frac{\text{True Positive}_i}{\text{True Positive}_i + \text{False Positive}_i} \tag{2}$$

$$\text{Recall}_i = \frac{\text{True Positive}_i}{\text{True Positive}_i + \text{False Negative}_i} \tag{3}$$

where $i = A, B, C, D$ According to the WBP score.

A general measure of Precision and Recall is calculated as the arithmetic mean of the values obtained for each individual class. Finally, F1 Score is obtained as the harmonic mean of Precision and Recall:

$$\text{F1} = \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

V. Numerical Results

The validation of the proposed method for image compression and reconstruction has been conducted by performing automatic wound classification with a CNN and comparing the classifier performance obtained on original images and those obtained on reconstructed images after CS. The network can detect and automatically classify the wound with a certain percentage of confidence. An example of the CNN output is shown in Fig.2-a and -b.

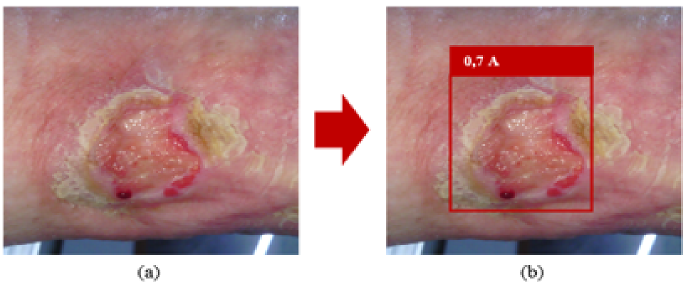

Fig. 2. (a) Input image. (b) Network detection and classification.

Regarding the original images, we achieved values of Precision equal to 0.812, Recall equal to 0.799 and F1 equal to 0.781. On the other hand, on the reconstructed images, the network obtained a Precision of 0.757, a Recall of 0.776. The CS technique implemented, based on the behavior model of a PCM array, appears promising, showing a reduction in classification performance of only 6.77% in terms of Precision, 2.88% in terms of Recall and 1.92% in terms of F1 Score compared to the original images.

VI. Conclusions

Phase-Change Memory is a promising technology with applications in non-volatile memory, computing elements for in-memory and neuromorphic computing and also as components of reconfigurable electronic circuits. In this work, we present the Compressed Sensing technique implemented by leveraging the conductive behavior model of a PCM array followed by an iterative decoding procedure to address the introduced non-linearity. The aim of this study is to find and validate a compression technique for skin ulcer images with a view to optimizing data transmission and storage in clinical settings, based on a device with memristive behavior. Additionally, CS has been extensively analyzed as a viable technique for image encryption, thus presenting itself as a method to enhance data privacy and security, which is of considerable relevance in healthcare [12]. The performance of lesion detection and automatic classification performed through CNN network on images before and after CS confirms the validity of the technique and therefore the use of images associated with a weight an order of magnitude lower than the original ones, while still retaining the relevant features for automatic classification of skin ulcers.

Acknowledgements

This project has been partially supported by the ECSEL Joint Undertaking (JU) under grant agreement No 101007321. The JU receives support from the European Union's Horizon 2020 research and innovation programme and France, Belgium, Czech Republic, Germany, Italy, Sweden, Switzerland, Turkey.

References

[1] J. Secco, E. Spinazzola, M. Pittarello, E. B. Ricci, and F. Pareschi, "Clinically validated classification of chronic woundsmethod with memristor-based cellular neural network," 2024.

[2] A. Antolini, C. Paolino, F. Zavalloni, A. Lico, E. F. Scarselli, M. Mangia, F. Pareschi, G. Setti, R. Rovatti, M. L. Torres *et al.*, "Combined hw/sw drift and variability mitigation for pcm-based analog in-memory computing for neural network applications," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 13, no. 1, pp. 395–407, 2023.

[3] Z. Sun, G. Pedretti, E. Ambrosi, A. Bricalli, W. Wang, and D. Ielmini, "Solving matrix equations in one step with cross-point resistive arrays," *Proceedings of the National Academy of Sciences*, vol. 116, no. 10, pp. 4123–4128, 2019.

[4] M. Le Gallo and A. Sebastian, "An overview of phase-change memory device physics," *Journal of Physics D: Applied Physics*, vol. 53, no. 21, p. 213002, 2020.

[5] M. Orlowski, J. Secco, and F. Corinto, "Chua's constitutive memristor relations for physical phenomena at metal–oxide interfaces," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 5, no. 2, pp. 143–152, 2015.

[6] L. O. Chua and S. M. Kang, "Memristive devices and systems," *Proceedings of the IEEE*, vol. 64, no. 2, pp. 209–223, 1976.

[7] F. Marrone, J. Secco, B. Kersting, M. Le Gallo, F. Corinto, A. Sebastian, and L. O. Chua, "Experimental validation of state equations and dynamic route maps for phase change memristive devices," *Scientific Reports*, vol. 12, no. 1, p. 6488, 2022.

[8] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on information theory*, vol. 52, no. 2, pp. 489–509, 2006.
[9] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing mri," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 72–82, 2008.
[10] T. Blumensath, "Compressed sensing with nonlinear observations and related nonlinear optimization problems," *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3466–3474, 2013.
[11] H.-C. Chen, H. Kung, and M. Comiter, "Nonlinear compressive sensing for distorted measurements and application to improving efficiency of power amplifiers," in *2017 IEEE international conference on communications (ICC)*. IEEE, 2017, pp. 1–7.
[12] Z. Zhang, N. Zhou, B. Sun, S. Banerjee, and J. Mou, "Multimedia healthcare cloud personal archives security system based on compressed sensing and multi-image encryption," *Journal of the Franklin Institute*, vol. 361, no. 8, p. 106844, 2024.