34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Spatiotemporal Compression for Efficient Storage and Transmission of High-Resolution Electrocorticography Data*

Taehoon Kim[1], N. Sertac Artan[1], Jonathan Viventi[1,2], H. Jonathan Chao[1]

***Abstract*— High-resolution Electrocorticography (HR-ECoG) has emerged as a key strategic technology for recording localized neural activity with high temporal and spatial resolution with potential applications in brain-computer interfaces (BCI), and seizure detection for epilepsy. However, HR-ECoG has 400 times the resolution of conventional ECoG, making it a challenge to process, transmit and store the HR-ECoG data. Therefore, simple and efficient compression algorithms are vital for the feasibility of implantable wireless medical devices for HR-ECoG recordings. In this paper, following the observation that HR-ECoG signals have both high spatial and temporal correlations similar to video/image signals, various compression methods suitable for video/image– compression based on motion estimation, discrete cosine transform (DCT) and discrete wavelet transform (DWT)– are investigated for compressing HR-ECoG data. We first simplify these methods to satisfy the low-power requirements for implantable devices. Then, we demonstrate that spatiotemporal compression methods produce up to 46% more data reduction on HR-ECoG data than compression methods using only spatial compression do. We further show that this data reduction can be achieved with low hardware complexity. In particular, among the methods investigated, spatiotemporal compression using DCT-based methods provide the best trade-off between hardware complexity and compression performance, and thus we conclude that DCT-based compression is a promising solution for ultralow-power implantable devices for HR-ECoG.**

I. INTRODUCTION

Recently, flexible electrode arrays (Fig. 1) were introduced for recording electrocorticogram signals at high resolution (HR-ECoG) [1].

HR-ECoG has a high temporal-domain correlation similar to that of the conventional ECoG. However, unlike conventional ECoG signals, HR-ECoG signals also have spatial-domain correlation, due to the close proximity (500-μm spacing) of the neighboring electrodes. This high spatiotemporal correlation allows for the observation of brain activity at an unprecedentedly finer granularity. However, to accommodate such higher resolution provided by HR-ECoG, a larger volume of data than that needed for conventional ECoG must be generated, transmitted and stored. Current HR-ECoG arrays already have 360 electrodes at 400 times higher resolution than conventional 64-channel ECoG arrays, with the potential of scaling in the future to thousands of electrodes [1]. This increase in data volume challenges the feasibility of medical implants for HR-ECoG. Therefore, simple and efficient compression algorithms are required to realize practical medical implants for HR-ECoG.

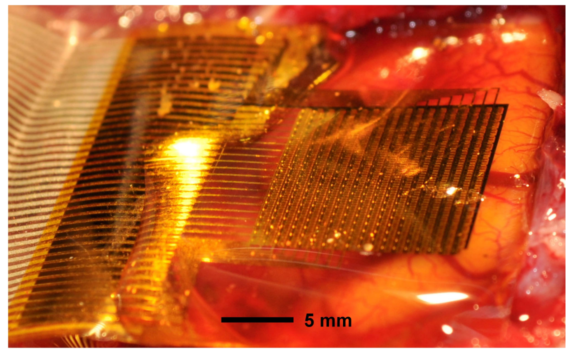


Fig. 1: Flexible, high-resolution multiplexed electrode array.

Various compression techniques for conventional ECoG or EEG have been proposed [2]–[4]. In [2], an H.264-based method is used for compressing multichannel recordings of evoked potentials, where multichannel neural video frames are generated such that a fixed number (S) of samples from channel 1 constitutes frame 1, and S samples from channel 2 constitutes frame 2, and so on. After N frames, where N is the number of channels, the next S samples from channel 1 constitute frame $N+1$, and the frame construction continues from channel 2 in a similar manner. This channel-by-channel frame structure requires additional preprocessing and buffer memory for generating each frame, since enough samples from each channel should be accumulated to constitute even a single frame. The channel-by-channel frame structure is a result of the the low spatial correlation of the conventional ECoG data, but it also complicates the real-time, low-power implementation of this algorithm for the conventional ECoG.

A 2D-DWT-based method is used for EEG compression in [3], using Daubechies 9/7 floating-point-filter coefficient and arithmetic encoder; providing superior compression performance, though with higher computational complexity. The authors in [4] focus on time-domain 1D-DWT with channel-by-channel multiplexing to eliminate temporal redundancy since the conventional ECoG has high temporal resolution but not necessarily high spatial resolution. Yet, time-domain 1D-DWT also needs a channel-by-channel frame structure, and thus it requires a certain amount of data to be saved in memory. Due to the necessity of increasing memory as the number of channels increase, the scalability of these methods

*Manuscript received March 31, 2012. This work was partially supported by a grant from Bonstar International Limited and a Taking Flight Award from Citizens United for Research in Epilepsy.

[1]T. Kim (phone: 1-718-260-3598; e-mail: tkim03@students.poly.edu), N. S. Artan (e-mail: artan@poly.edu), J. Viventi (e-mail: jviventi@poly.edu) and H. J. Chao (e-mail: chao@poly.edu) are with the Department of Electrical and Computer Engineering at the Polytechnic Institute of New York University, Brooklyn, NY 11201 USA.

[2] J. Viventi is also with the Center for Neural Science at New York University, New York, NY 10003 USA.

for a large number of channels is limited. HR-ECoG, on the other hand, enables both spatial and temporal compression so that we can get a better compression performance while operating at a low clock frequency using smaller memory.

In this paper, we compare and simplify several video/image compression methods; based on H.264, DCT and DWT; for low power implantable devices using HR-ECoG data. The proposed DCT-based spatiotemporal compression methods show up to 46% more data reduction than using only spatial compression.

The rest of this paper is organized as follows: In Section II, we present the properties of HR-ECoG signals. In Section III, several compression algorithms using image/video compression methods are introduced. In Section IV, performance results of these methods for HR-ECoG and their complexity for real-time processing are provided. In Section V, the effect of reconstructed data and future works are discussed. Section VI concludes the paper.

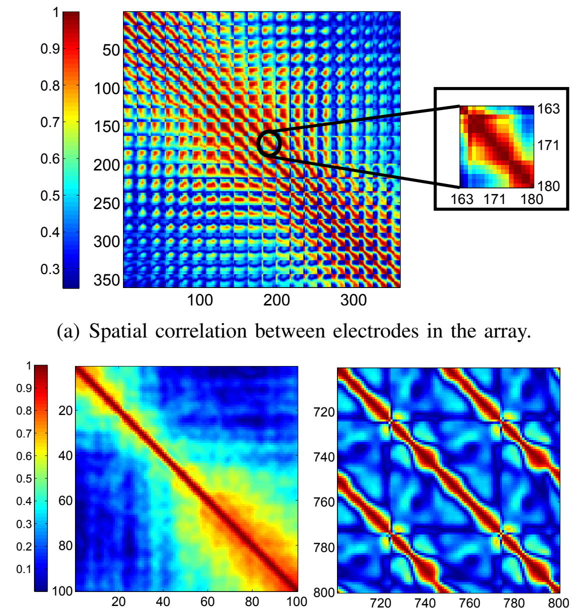
II. Properties of High-Resolution ECoG

The current HR-ECoG arrays consist of 360 electrodes organized in an 18×20 grid, in which each channel is sampled at 278 Hz. HR-ECoG has higher spatial resolution than the conventional ECoG due to the close proximity of the electrodes in HR-ECoG. This higher resolution leads to a high spatial correlation between neighboring electrodes. An example of high spatial correlation between neighboring electrodes can be seen in Fig. 2 (a), which shows the correlation between electrodes on the data recorded during electrographic seizures of a cat brain *in vivo* (*e.g.,* the inset in Fig. 2 shows the high correlation between electrodes in the same row ranging from electrode 163 to 180.)[1]. The correlation is lower as electrode pairs farther apart are considered, as illustrated in Fig. 2 (a). Therefore, HR-ECoG signals have a more prevalent spatial correlation similar to that of the video/image signals. As a result, HR-ECoG frames can be constructed either per electrode or per time instance, unlike the conventional ECoG, which only favors per channel frame construction as discussed in Section I. Fig. 2 (b) shows the temporal correlation for frames in two different ranges: The left graph shows the frame range 1 to 100 and the right graph shows the frame range 701 to 800. Note that although the correlation in between frames (*i.e.,* temporal correlation) is high for this seizure data, it is not as high as spatial correlation as shown in Fig. 2. Yet, as shown in Section IV, this temporal correlation will boost the compression performance if used together with spatial compression.

After the analog-to-digital conversion, a set of 18×20 samples acquired in a single time slot (*i.e.,* one sample from each channel totaling 360 samples) can be used directly as a frame without any modification (*i.e.,* the frame organization is the same as the spatial organization of the electrodes, similar to the relation between a scene and a video frame).

[1]Here, correlation values are scaled between 0 and 1, where a value of 1 shows the highest correlation between two electrodes. A higher correlation indicates redundancy in the data and thus offers better compression.

This frame structure, which relies only on samples acquired at the current time, eliminates the necessity of preprocessing and buffer memory. These necessities, as pointed out in Section I, limit the practicality of the conventional ECoG [2], but not that of the HR-ECoG.



(a) Spatial correlation between electrodes in the array.

(b) Temporal correlation between frames.

Fig. 2: Correlation of HR-ECoG

III. Proposed Compression Algorithms

In this section, we introduce common video/image compression methods and describe how to apply them to compress the HR-ECoG signals. There are 3 representative compression methods in video/image compression: H.264, JPEG and JPEG2K. H.264 is a video compression standard using motion estimation [5]. The JPEG image compression standard [6] adopts DCT as the core algorithm, while JPEG2K image compression standard [7] uses DWT as the core algorithm. These algorithms usually require high computation power and large storage, and thus are not suitable for implantable devices. In this paper, we simplify these algorithms to satisfy the low power requirement for implantable devices.

A. H.264-Based Compression Methods

H.264 is a widely deployed video coding standard, which supports variable macroblock (MB) sizes such as 16×16, 16×8, 8×8 and 4×4 [5]. Using variable MB provides a compression gain up to $2.0\times$ over previous standard such as H.263+ and MPEG-4 Part 2 [8]. However, this variable block processing complicates the motion estimation step of H.264, with respect to previous video coding standards. Especially, the integer-pel motion estimation (IPME) takes almost 80% of the computation and memory access in H.264 [9].

To better fit the low-power requirements of implantable devices, in this paper, we simplify the motion estimation

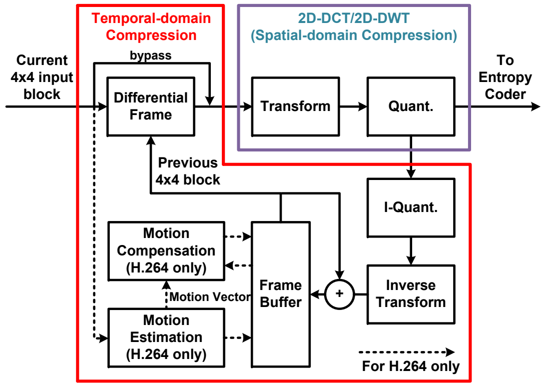


Fig. 3: Block diagram for the proposed HR-ECoG compression architecture.

(ME) operation of the H.264 to consume less power. The proposed ME block, shown in Fig. 3, uses only a 4×4 MB IPME with a horizontal and vertical ± 8 search range. Every N frames interval, called I-frame, is used as a reference frame to predict the next frame. In I-frame, each 4×4 input block is transformed using 2D-DCT. After the DCT, the coefficients are quantized, zig-zag ordered and entropy coded. The rest of the frames (the P-frames) use the inter-mode coding. In inter-mode coding, the ME block seeks the best matching block in the previous frame and finds its motion vector (MV). Then the residual data (the difference between the previous block and the current block) is passed through DCT and quantization blocks in the same manner as that in intra-mode coding. In motion compensation, previous frames are reconstructed using their MVs.

B. DCT-Based Compression Methods

The DCT has been widely deployed in popular image and video coding standards such as JPEG and MPEG. In JPEG, 8×8 2D-DCT is used for converting an image into a frequency representation [6]. However, 8×8 2D-DCT has high hardware complexity. Here we propose to use a 4×4 Integer DCT for HR-ECoG, due firstly to the high correlation within the 4×4 neighboring array, and secondly to the fact that 4×4 integer DCT requires only adders and shifters but no multiplier [10], thus significantly reducing the power requirement. For the 2D-DCT, we use only spatial compression as in JPEG. The 4×4 input is directly 2D-DCT transformed, quantized, zig-zag ordered, and entropy coded.

If both spatial and temporal compression are to use DCT, then the Differential Frame 2D-DCT (DF-2D-DCT) can be used. An optional DF-2D-DCT module can be used as shown in Fig. 3 to calculate the difference between the previous 4×4 input block and the current 4×4 input block to eliminate the temporal redundancy. After the data is processed in the DF module, the residual data is first sent to the 4×4 DCT block to remove spatial redundancy and then passed to the quantization module in the same manner as in 2D-DCT.

C. DWT-Based Compression Methods

The DWT has become an important method for a variety of signal processing applications including image compression. In JPEG2K, the floating-point Daubechies 9/7 filter and arithmetic encoder are adopted for higher compression [7]. In this paper, we propose to use the 2-level DWT with the biorthogonal spline 5/3 filter because the complexity of DWT depends on the length of the filter coefficients and the number of levels [11]. Also, the Huffman encoder is substituted for arithmetic encoder, because Huffman coding implementation is less complex thus power-efficient.

Note that the difference between DWT-based and DCT-based methods is only in the transformation block in Fig. 3. In the 2D-DWT method, we use only spatial compression like that with JPEG2K. Input data are directly 2D-DWT transformed, quantized, reordered and entropy coded.

IV. Results

To evaluate the performance of the compression algorithms for HR-ECoG, we use two separate HR-ECoG recordings [1] from 360 channel HR-ECoG arrays: (1) 550 seconds of neural recording data for evoked potentials. (2) 6 seconds of neural recording data during an electrographic seizure.

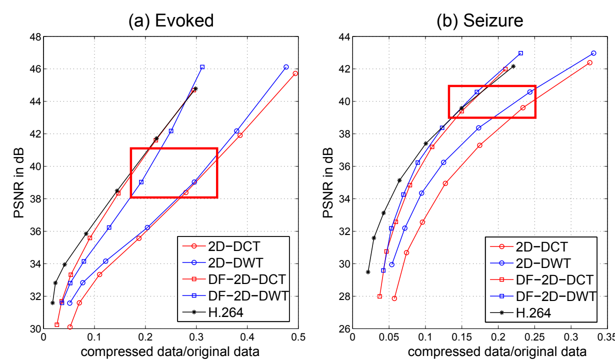


Fig. 4: Comparison of the performance of the compression algorithms for high-resolution ECoG.

Fig. 4 shows the peak signal-to-noise ratio (PSNR) for various methods for different compression ratios. For both recordings, spatiotemporal compression methods consistently show better compression performance for a wide range of compression ratios. For evoked potentials, around a PSNR of 40 dB, spatiotemporal compression methods provide up to 46% more data reduction than only spatial compression.

TABLE I: The minimum required clock frequency for real-time compression (column 3) and the memory usage (column 4) of the proposed compression algorithms.

Algorithm	Correlation	f_{clk} [kHz]	Memory [kByte]
2D-DCT	Spatial	112	0.8
2D-DWT	Spatial	186	1.2
DF-2D-DCT	Spatiotemporal	224	1.2
DF-2D-DWT	Spatiotemporal	372	1.6
H.264	Spatiotemporal	7,100	1.6

It is generally accepted that when the PSNR is above $40\ dB$, the reconstructed image is virtually indistinguishable from the original image. Based on the two ECoG recordings used in this paper, at PSNR values at or above $40\ dB$, there is no loss of neural oscillations, and for the data presented it is acceptable for visual analysis as demonstrated in Fig. 5.

Table I shows the complexity of the proposed algorithms. To achieve real-time processing, each frame should be compressed within 3.6 ms (a sampling of 360 channels at 278 Hz takes around $T = 3.6$ ms). The clock frequency, F_{clk}, for real-time processing is estimated from the memory access rate assuming a 16-bit data bus and a memory throughput of one sample per clock frequency (*i.e.,* SRAM). Motion estimation, the most time consuming task of H.264, requires $T_r = 16$ clock cycles to read each 4×4 MB and a 18×20 array contains $N = 25$ MBs. Therefore, the number of clock cycles required for motion estimation to process one frame using ± 8 search range is $N \times T_r \times 8 \times 8 \approx 25,600$ cycles. Therefore, $f_{clk} \geq 25,600/T = 7.12\ MHz$. In the DCT case, there are 25 4×4 blocks per frame, and no search is needed and so $f_{clk} \geq 16 \times 25/T = 112\ KHz$. 2D-DWT processes both the horizontal and vertical axes twice. It requires 12 cycles to read one horizontal line and for both horizontal and vertical processing $12 \times 20 \times 2 = 480$ cycles are required per frame. Since we use 2-level 2D-DWT, we need an additional 7 cycles $\times$ 10 lines $\times\ 2 = 140$ cycles for 2-level 2D-DWT, resulting in an $f_{clk} = 173\ kHz$.

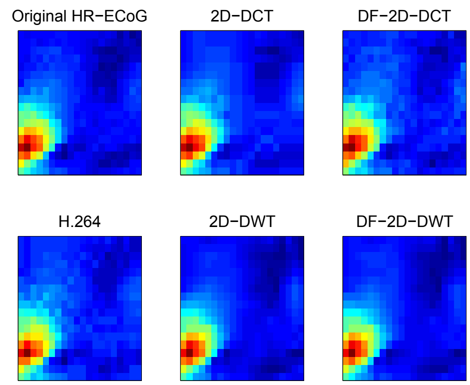


Fig. 5: Reconstructed two-dimensional data using different compression methods @PSNR = 40 dB.

V. Discussion

The effect of the proposed compression methods was evaluated by calculating the average correlation coefficient r between original HR-ECoG data and reconstructed HR-ECoG data. We calculated the average correlation coefficient for each of the reconstructed HR-ECoG data. The average r is above 0.97 at PSNR $40\ dB$ for each of the proposed compression algorithms. Based on these results, we can expect that reconstructed data above PSNR $40\ dB$ can be used for prediction of visual evoked responses as shown in [1] without performance degradation. We will investigate the effect of compression on the accuracy of predicted visual evoked responses in future work. Due to the complexity of H.264, it is unfavorable for low-power implant devices. DF-2D-DCT has similar compression performance as DF-2D-DWT, but it has a lower clock frequency and memory requirement. In addition, 2D-DWT consumes about two times more power than 2D-DCT [12].

In summary, the DF-2D-DCT algorithm is the most suitable for ultra-low-power implantable devices for HR-ECoG: compressing the data volume to 16% for electrographic seizure, 18% for evoked potentials, while achieving a PSNR of $40\ dB$ after reconstruction. For the seizure data, since a PSNR of around 28 dB is adequate, HR-ECoG data can be compressed to only 4% of its original size.

VI. Conclusion

HR-ECoG has tremendous potential for many research and clinical applications. In this paper, based on the similarities between HR-ECoG and video, we present various compression methods for HR-ECoG. We show that better compression performance is obtained when using spatial-temporal compression rather than the spatial compression with reasonable hardware complexity. In particular, we showed that the DF-2D-DCT method is a very simple and efficient compression algorithm that provides the best trade-off between compression ratio and power consumption when using implantable devices for HR-ECoG.

References

[1] J. Viventi *et al.*, "Flexible, foldable, actively multiplexed, high-density electrode array for mapping brain activity in vivo," *Nature Neuroscience*, vol. 14, no. 4, pp. 1599 – 1605, 2011.

[2] C. H. Chung, L.-G. Chen, Y.-C. Kao, and F.-S. Jaw, "Multichannel evoked neural signal compression using advanced video compression algorithm," in *4th International IEEE/EMBS Conference on Neural Engineering (NER 2009)*, May 2009.

[3] G. Higgins, B. McGinley, E. Jones, and M. Glavin, "Efficient EEG compression using JPEG2000 with coefficient thresholding," in *Signals and Systems Conference (ISSC 2010), IET Irish*, Jun. 2010.

[4] K. Oweiss, A. Mason, Y. Suhail, A. Kamboh, and K. Thomson, "A scalable wavelet transform vlsi architecture for real-time signal processing in high-density intra-cortical implants," *IEEE Trans. Circuits and Systems I*, vol. 54, no. 6, pp. 1266 –1278, Jun. 2007.

[5] D. Marpe, T. Wiegand, and G. Sullivan, "The H.264/MPEG4 advanced video coding standard and its applications," *IEEE Communications Magazine*, vol. 44, no. 8, pp. 134 –143, Aug. 2006.

[6] G. Wallace, "The JPEG still picture compression standard," *IEEE Trans. Consumer Electronics*, vol. 38, no. 1, Feb. 1992.

[7] A. Skodras, C. Christopoulos, and T. Ebrahimi, "The JPEG 2000 still image compression standard," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 36 –58, Sep. 2001.

[8] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G. Sullivan, "Rate-constrained coder control and comparison of video coding standards," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 688 – 703, 2003.

[9] T.-C. Chen, S.-Y. Chien, Y.-W. Huang, C.-H. Tsai, C.-Y. Chen, T.-W. Chen, and L.-G. Chen, "Analysis and architecture design of an HDTV720p 30 frames/s H.264/AVC encoder," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 16, no. 6, Jun. 2006.

[10] H. Malvar, A. Hallapuro, M. Karczewicz, and L. Kerofsky, "Low-complexity transform and quantization in H.264/AVC," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 13, no. 7, 2003.

[11] B. Usevitch, "A tutorial on modern lossy wavelet image compression: foundations of JPEG 2000," *Signal Processing Magazine, IEEE*, vol. 18, no. 5, pp. 22 –35, sep 2001.

[12] M. Lanuzza, M. Margala, and P. Corsonello, "Cost-effective low-power processor-in-memory-based reconfigurable datapath for multimedia applications," in *Proceedings of the International Symposium on Low Power Electronics and Design.*, Aug. 2005, pp. 161 – 166.