

# A simple and efficient algorithm operating with linear time for MCEEG data compression

Geevarghese Titus<sup>1,2</sup> · M. S. Sudhakar<sup>1</sup>

Received: 3 January 2017 / Accepted: 23 July 2017 / Published online: 31 July 2017  
© Australasian College of Physical Scientists and Engineers in Medicine 2017

**Abstract** Popularisation of electroencephalograph (EEG) signals in diversified fields have increased the need for devices capable of operating at lower power and storage requirements. This has led to a great deal of research in data compression, that can address (a) low latency in the coding of the signal, (b) reduced hardware and software dependencies, (c) quantify the system anomalies, and (d) effectively reconstruct the compressed signal. This paper proposes a computationally simple and novel coding scheme named spatial pseudo codec (SPC), to achieve lossy to near lossless compression of multichannel EEG (MCEEG). In the proposed system, MCEEG signals are initially normalized, followed by two parallel processes: one operating on integer part and the other, on fractional part of the normalized data. The redundancies in integer part are exploited using spatial domain encoder, and the fractional part is coded as pseudo integers. The proposed method has been tested on a wide range of databases having variable sampling rates and resolutions. Results indicate that the algorithm has a good recovery performance with an average percentage root mean square deviation (PRD) of 2.72 for an average compression ratio (CR) of 3.16. Furthermore, the algorithm has a complexity of only  $O(n)$  with an average encoding and decoding time per sample of 0.3 ms and 0.04 ms

respectively. The performance of the algorithm is comparable with recent methods like fast discrete cosine transform (fDCT) and tensor decomposition methods. The results validated the feasibility of the proposed compression scheme for practical MCEEG recording, archiving and brain computer interfacing systems.

**Keywords** Inverse logarithmic translation (ILT) transform · Integer fraction coder (IFC) · Integer fraction decoder (IFD) · Pseudo integer (PI) · Translated logarithmic (TL) transform

## Introduction

Electroencephalograph (EEG) [1] is the recording of electrical activity that is used to scrutinize the neurophysiological properties of the brain [2], to diagnose brain disorders like epilepsy, sleep disorders and in various brain computer interfaces (BCI) applications like mind-controlled gaming, neuromarketing and social interaction analysis [3]. Such applications require robust portable EEG recorders like NeuroSky MindWave, Emotiv Neuroheadset and Imec EEG Headset, that record single or multichannel EEG data for an extended period (hours, days, or potentially, even months) at varying sampling frequencies and resolutions. This continuous recording of EEG can quickly generate large amounts of data, causing a serious concern in terms of power and storage requirements. To enable such portable devices to work effectively, research on compression algorithms is of prime importance.

EEG compression algorithms [4–7] exploit various spatial and transform domain properties of the signal—like context, inter/intra-channel correlation, energy compaction, residual energy and power spectrum. There

✉ M. S. Sudhakar  
sudhakar.ms@vit.ac.in

Geevarghese Titus  
geevarghese.titus2013@vit.ac.in;  
geevarghesetitius@amaljyothi.ac.in

<sup>1</sup> School of Electronics Engineering, VIT University, Vellore, India

<sup>2</sup> Department of Electronics and Communication, Amal Jyothi College of Engineering, Kottayam, Kerala, India

exists a trade-off between attainable compression and the degree of distortion in the reconstructed signal that is explored in all lossy and near lossless compression algorithms.

Previous methodologies employed in compression include transforms like Karhunen-Loeve [8], Wavelet [9–11]; dimensionality reduction techniques like principal component analysis (PCA) [12], independent component analysis (ICA) [13, 14]; compressive sensing techniques [15, 16]. Compression can be further enhanced by lossless or lossy coding techniques like arithmetic coding [17], set partition in hierarchical trees (SPIHT) [18]. Near lossless compression can be realized by employing error predictors like neural network and linear sequential predictors [19, 20].

Recent works in MCEEG including tensor decomposition using parallel factor decomposition (PARAFAC), singular value decomposition (SVD) [21], wavelet image transform followed by volumetric coding [22] guarantee the maximum possible error and exploit inter/intra-channel correlation; hybrid system of PCA, FastICA with SPIHT coding [23] takes advantages of both PCA and FastICA for reducing the dimensionality and fDCT [24] exploit transform domain properties to reduce the complexity of the system. To achieve compression, most of these methods rely on computationally intensive progressive operations, limiting its use in real time scenarios. Moreover, none of these methods normalize the data, so that contribution of all samples including outliers can be considered. These outliers need not be random errors but triggered motor or imagery responses to visual, audio or other stimuli. So, there is a need to explore possible hardware-friendly schemes that operate with lower latency giving good reconstruction performance and quantifiable distortion, for normalization and compression.

In this paper, a novel and computationally simple system, SPC is proposed. The scheme demonstrates how the concept of data normalization and representation can be utilized in MCEEG compression. First, an architecture that normalizes the MCEEG data using translation logarithmic (TL) transform is proposed, followed by an integer-fractional coder (IFC) that encodes the data, resulting in data compression. Finally, the effectiveness and robustness of the scheme is validated, using distortion indicators like PRD and PSNR.

Systematization of the paper is as follows: Section 2 discusses the methodology employed in the MCEEG compression, followed by a discussion on evaluation procedures in Sect. 3. Results and discussions are elaborated in Sect. 4. Complexity analysis of the algorithm is explored in Sect. 5 and the paper is concluded in Sect. 6.

## Methodology for compression

MCEEG signals are random, bipolar and skewed in nature. These signals can be acquired at varying sampling rates ranging from 250 to 2000 Hz, and digitized at 12/16/24 bit resolution. These signals are processed in IEEE single or double precision floating point format [25–27] by the computing systems. The proposed near lossless compression is a two-stage process where the MCEEG signal is normalized using TL block which is subsequently coded in the IFC block. The processes are detailed in the subsequent subsections and illustrated in Fig 1.

### Encoding process

The encoder of the SPC system is illustrated in Fig 1a. The first step in encoding is the arrangement of the MCEEG data to a 2D structure as shown in (1), where  $M$  and  $N$  correspond to the number of channels and samples per channel respectively.

$$x = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1N} \\ X_{21} & X_{22} & \dots & X_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ X_{M1} & X_{M2} & \dots & X_{MN} \end{bmatrix} \quad (1)$$

To exploit the contribution of all samples and make them suitable for the subsequent processing, the data is pre-processed by the TL transform, which performs two operations: translation and logarithm transformation. To ascertain that the resulting normalized samples are real and positive, the MCEEG samples are translated by a factor  $f$ , that depends on the gain of the employed MCEEG recorders.

$$g(x) = T(x + f) \quad (2)$$

Natural logarithm is deployed in the algorithm for normalization; whilst other bases do not contribute to any significant improvement in compression and signal recovery are not discussed here.

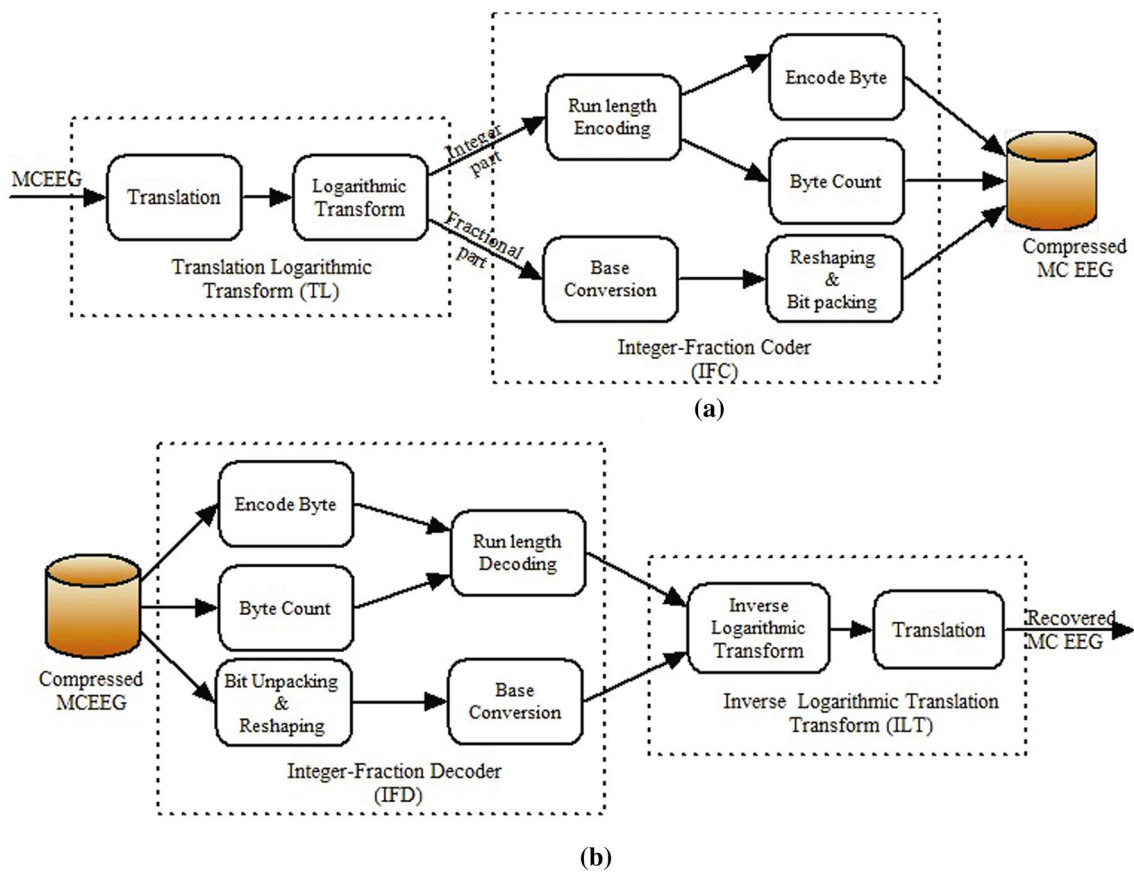
$$h(x) = \ln g(x) \quad (3)$$

The normalized data  $h(x)$  is split into integer  $I_r(x)$  and fractional  $I_f(x)$  parts using (4) and (5).

$$I_r(x) = \lfloor h(x) \rfloor \quad (4)$$

$$I_f(x) = h(x) - I_r(x) \quad (5)$$

$I_r(x)$  and  $I_f(x)$  are encoded separately using the IFC.  $I_r(x)$  can be encoded using any spatial coding schemes that exploit the redundancies introduced by the TL transform. In the current architecture, run length encoding (RLE) [28] is employed because of its simplicity in



**Fig. 1** Proposed SPC system (a) encoder section (b) decoder section

implementation. The encoded data are represented as  $(D_1, C_1)(D_2, C_2) \dots (D_i, C_i) \dots (D_n, C_n)$ , where  $D_i$  and  $C_i$  are the  $i$ th distinct integer and their occurrence (count) respectively. As arithmetic coders and Huffman coders introduce complexity in the system with only a marginal improvement in compression, they are not considered suitable in the proposed architecture.

Subsequently,  $I_f(x)$  can be encoded to an equivalent representation with error deviation depending on the bit depth  $d$  of the base converter. First, the converter converts  $I_f(x)$  to its equivalent binary stream  $b$  using (6), which is the generalized relation for converting fractions to or from any base.

$$b = a_0m^{-1} + a_1m^{-2} + a_2m^{-3} + \dots + a_{k+1}m^{-k} \tag{6}$$

where  $m$  is the base,  $k$  is the resolution and  $a$  takes values  $0, 1, \dots, m - 1$ . For example, for converting the fraction to binary,  $m$  is 2, and  $a$  takes values 0 or 1.

Data representations for faster encoding and decoding [29] include methods like variable byte, byte-oriented encoding, binary packing, binary packing with variable-length blocks and patched coding. Such representations are

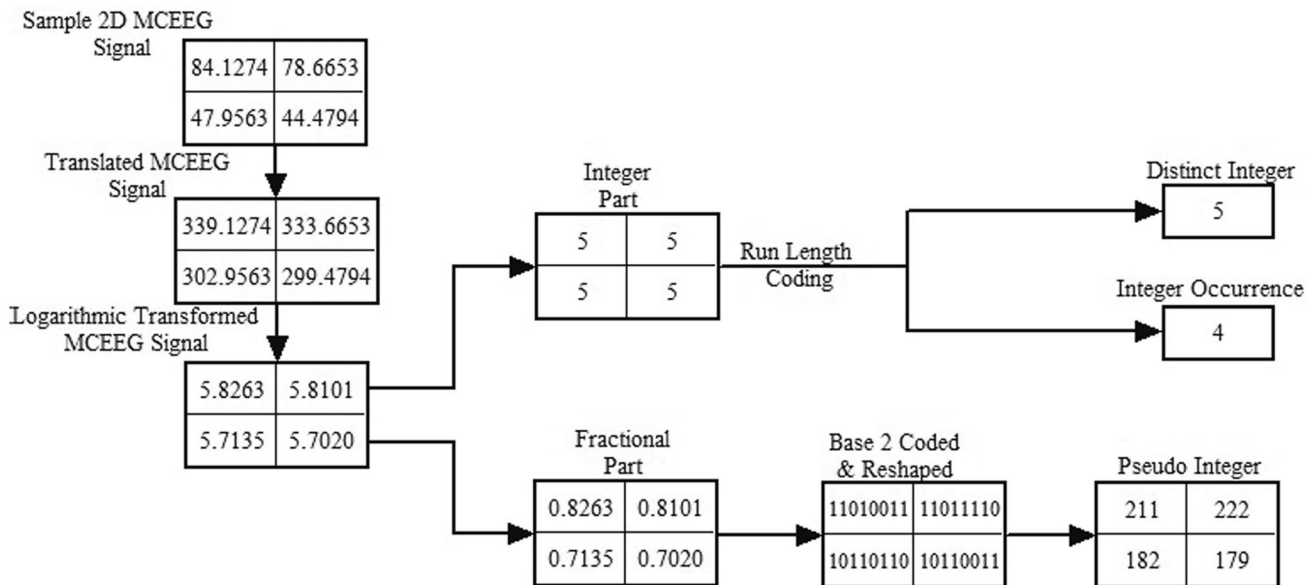
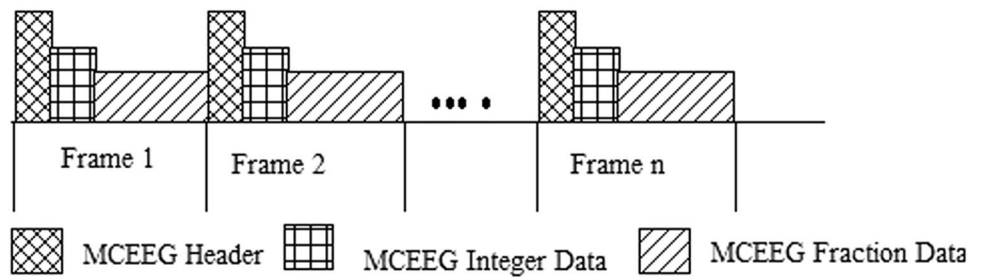
also useful in size reduction and faster access. A similar packing strategy of the binary stream has been employed in this architecture.

For data representation in the proposed algorithm, the binary stream  $b$  from base converter is reshaped and packed into groups of eight bits to form an integer equivalent representation of the fractional data called PI. This representation supplemented with spatial coding technique, helps the algorithm to achieve compression. The coded MCEEG is stored as frames as depicted in Fig 2, having three fields MCEEG header, integer data and fraction data. Figure 3 illustrates the encoding of four samples of 2D MCEEG signals being encoded and represented by distinct integer, integer occurrence and pseudo integers.

**Decoding process**

The decoder of the SPC system is illustrated in Fig 1b. For decoding the compressed MCEEG consisting of the RLE encoded data and PI, is first processed in the integer fractional

**Fig. 2** Framing process in MCEEG compression



**Fig. 3** Sample MCEEG encoding process

decoder (IFD). The RLE data is unpacked by interpolating the coded integer  $D_i$  by the parameter  $C_i$  to obtain  $I_r(x')$ . Subsequently, the PI are binarized and reshaped according to the bit depth value used for encoding, which is retrieved from MCEEG Header. The reshaped data is converted to fractional base using (6) to obtain  $I_f(x')$ . Both processes can be performed in parallel hence, reducing the decoding time.  $I_f(x')$  is added with  $I_r(x')$  resulting in the reconstructed normalized signal  $I(x')$ .

$$I(x') = I_r(x') + I_f(x') \tag{7}$$

Inverse transformation of  $I(x')$  is performed in the ILT block where an exponential operation defined by (8) is performed initially.

$$h(x') = \exp I(x') \tag{8}$$

Subsequently the reconstructed MCEEG signal  $x'$  is obtained by translating  $h(x')$ , given by (9), where factor  $f$ , is the same as that used in the encoder.

$$x' = T(h(x') - f) \tag{9}$$

**Table 1** MCEEG data sets used for testing

Dataset	No. of channels	Sampling frequency (Hz)	Dataset source	Resolution (bits)
1	22	250	BCIC IV Set 2a	16
2	31	500	SCCN	16
3	64	256	UCI MLR	16
4	28	1000	BCIC II Set 4a	16
5	28	100	BCIC II Set 4b	16
6	64	240	BCIC III Set 2	16
7	64	1000	BCIC IV Set 1	16
8	64	80	Physionet	12

BCIC brain computer interface competitions, SCCN swartz center for computational neuroscience, UCI MLR UCI machine learning Repository

**Performance validations**

To evaluate the efficacy of the proposed scheme, few data sets have been chosen, labelled as data set 1–8, from each

of these standard databases illustrated in Table 1. The databases used are from Berlin BCI Competitions (II–IV) [30–32], Swartz Center for Computational Neuroscience (SCCN) [33] and PhysioNet [34]. The database is made up of MCEEG recording of different subjects performing various motor or imagery tasks, subjected to various constraints and stimuli. Data sets 1, 4 and 5 are processed in IEEE double precision floating point format whereas data sets 2, 3, 6 and 7 in IEEE single precision floating point.

### Performance measures

The efficacy of the proposed method is validated based on various quality indicators [35], like compression ratio (CR), maximum absolute error (MAE), percentage root mean square difference (PRD), root mean square error (RMSE) and peak signal to noise ratio (PSNR). As PRD alone does not convey the distortion in the reconstructed signal, other measures like PSNR, MAE, PAE and RMSE and their relations are also investigated.

### Results and discussion

The proposed architecture of MCEEG compression algorithm shown in Fig. 1 was evaluated using the data set discussed in Sect. 3. All channels of the EEG data sets were taken for simulation, though increasing or reducing the number of channels or samples per channel did not have any significant impact on the degree of compression. A visual illustration of the original, reconstructed and error signal at a bit depth of 5, is illustrated in Fig. 4. The bit depth of five can be considered as an optimal choice as it gave

almost similar and relatively good CR and distortion measures for all the data sets.

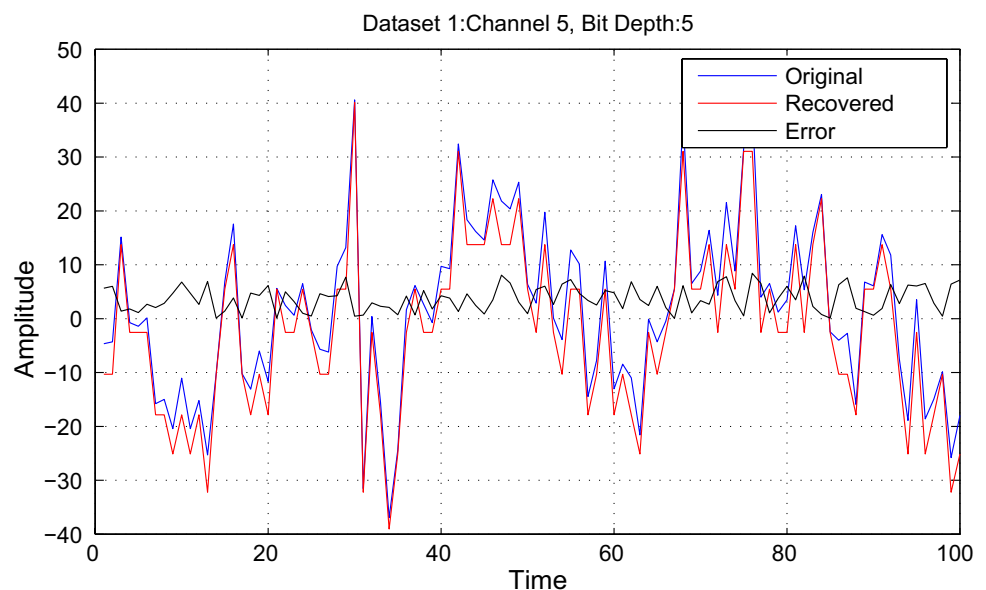
The logarithmic transformation normalized the data, and the process is fully reversible. The reconstructed signal quality was visually validated and quantified numerically using the least absolute error (LAE) distortion parameter. The maximum error in the reconstructed signal was in the order of  $10^{-7}$ – $10^{-9}$ . The CR and distortion performance for different bases of the logarithmic transform showed no significant variation hence, are not discussed further.

### Observation of the proposed algorithm

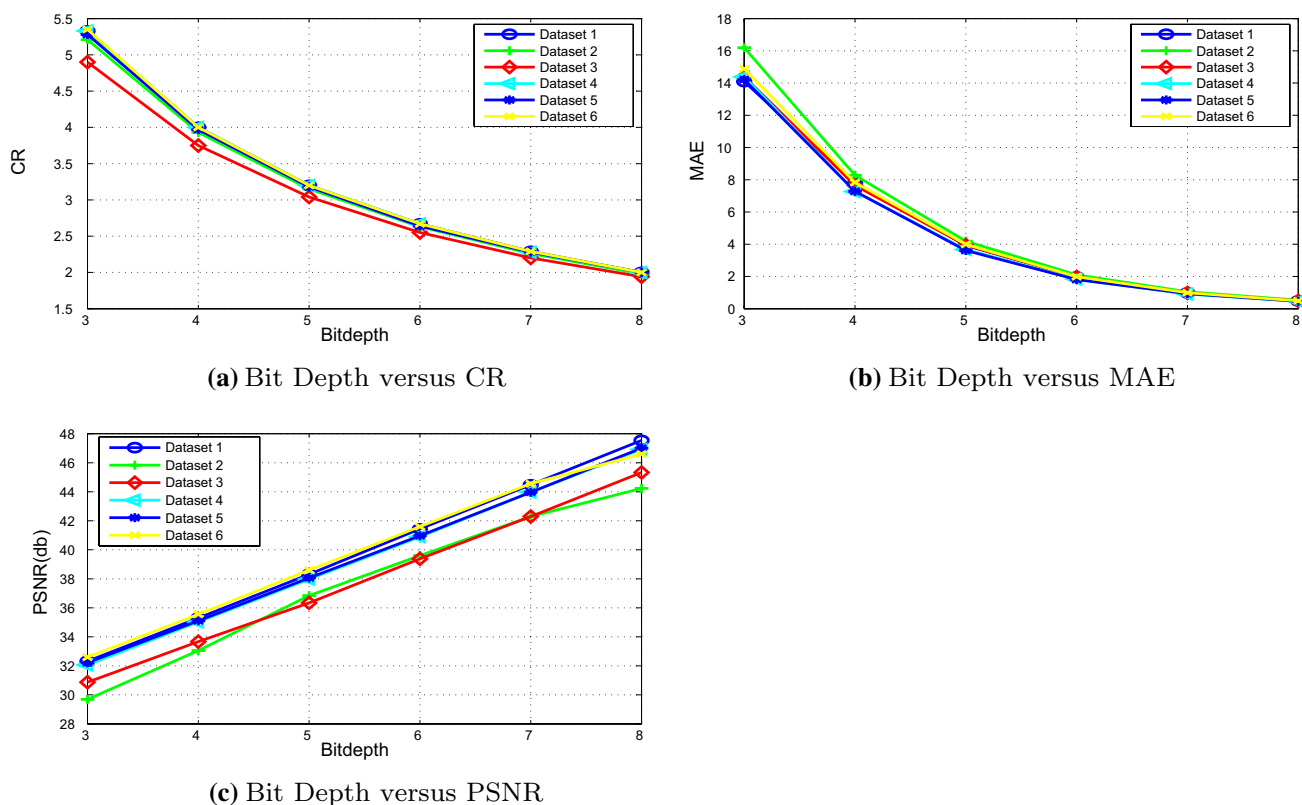
The performance of the proposed algorithm for different data sets are illustrated in Figs. 5, 6, 7. The bit depth  $d$  of base converter contributes largely to the compression as well as distortion in the recovered signal, as illustrated by Fig. 5a–c. The lower value of bit depth results in higher compression, with larger distortion in the recovered signal and vice versa. Hence, choice of optimal value of bit depth depends on the acceptable amount of distortion in the reconstructed signal. This value was found out with visual scoring along with numerical values of the distortion metrics PRD, RMSE, PAE, MAE and PSNR as elaborated in the subsequent paragraphs.

It was observed that at higher bit depths distortion is negligible and CR is comparatively less. For e.g., from Figs. 5f, 6a, d, CR at a bit depth of eight is nearly two, with PRD value very close to zero and PSNR above 44 db. The performance of the algorithm at different sampling frequencies (100, 1000 Hz) is almost the same as exemplified by distortion indicators corresponding to data sets 4 and 5 in the performance plots. Another

**Fig. 4** Superimposed original, reconstructed and residual error signals at an optimal bit depth of five







**Fig. 5** Performance illustration of proposed scheme of different MCEEG data sets for varying bit depth

observation is that the resolution of the original data does play a significant role in achievable CR. It can be concluded that higher CR can be achieved at higher resolutions of the ADCs and vice versa. This is a significant observation as currently available ADCs are operating at higher resolutions than those employed in this study.

Distortion performance indicators vary among the data sets at the same bit depth. PAE value indicates the largest difference between the original and reconstructed signal. MAE, on the other hand, indicates the mean error in the reconstructed signal. RMSE value, apart from being an error measure also serves as a performance indicator of the outliers. There was a four-fold increase in the RMSE value and a two-fold increase in the PAE and MAE values for a successive reduction in bit depth.

For a bit depth of five, the acceptable average CR of 3.16 occurs at average MAE value of 3.88. Moreover, from Fig. 7b, at the average value of MAE, the maximum variation of PAE ranges from 10 to 15. This provides an upper limit to the sample error at the current CR. Furthermore, average PSNR of 37.69 db and PRD of 2.72 are observed at the specified CR. Hence, based on the visual scoring and critical inference from Figs. 5, 6, 7, it can be concluded that the optimal bit depth across different data sets can be taken as five, though in some data sets much

lower bit depths can also be considered with similar distortion parameters.

Table 2 illustrates the encoding and decoding time of the proposed algorithm for different number of channels and samples per channel. Computations were performed on a PC with Intel Core2Duo processor operating at 1.8 GHz with 2 GB RAM. From the table, the average encoding and decoding times per sample can be computed, which are 0.3 and 0.04 ms respectively. This is a clear indicator that the algorithm is computationally simple and fast when compared to other progressive computation algorithms.

Use of public and common data sets are required for relative performance analysis of novel compression algorithms. In the proposed work, data sets 4, 7 and 8 were used for performance comparison as they were the ones used in recent compression algorithms [22, 24, 36]. A comparison is illustrated in Table 3, taking PRD and PSNR as reference quality indicators and the best CR measure is indicated in bold. The comparative results signify that the performance of the proposed method in terms of compression and distortion is comparable with recent work but with a much lower time complexity. The main highlight of the proposed algorithm is its simplicity in compressing and decompressing the MCEEG data [24] claims to be computationally light, but it is not superior to the proposed algorithm, as observed

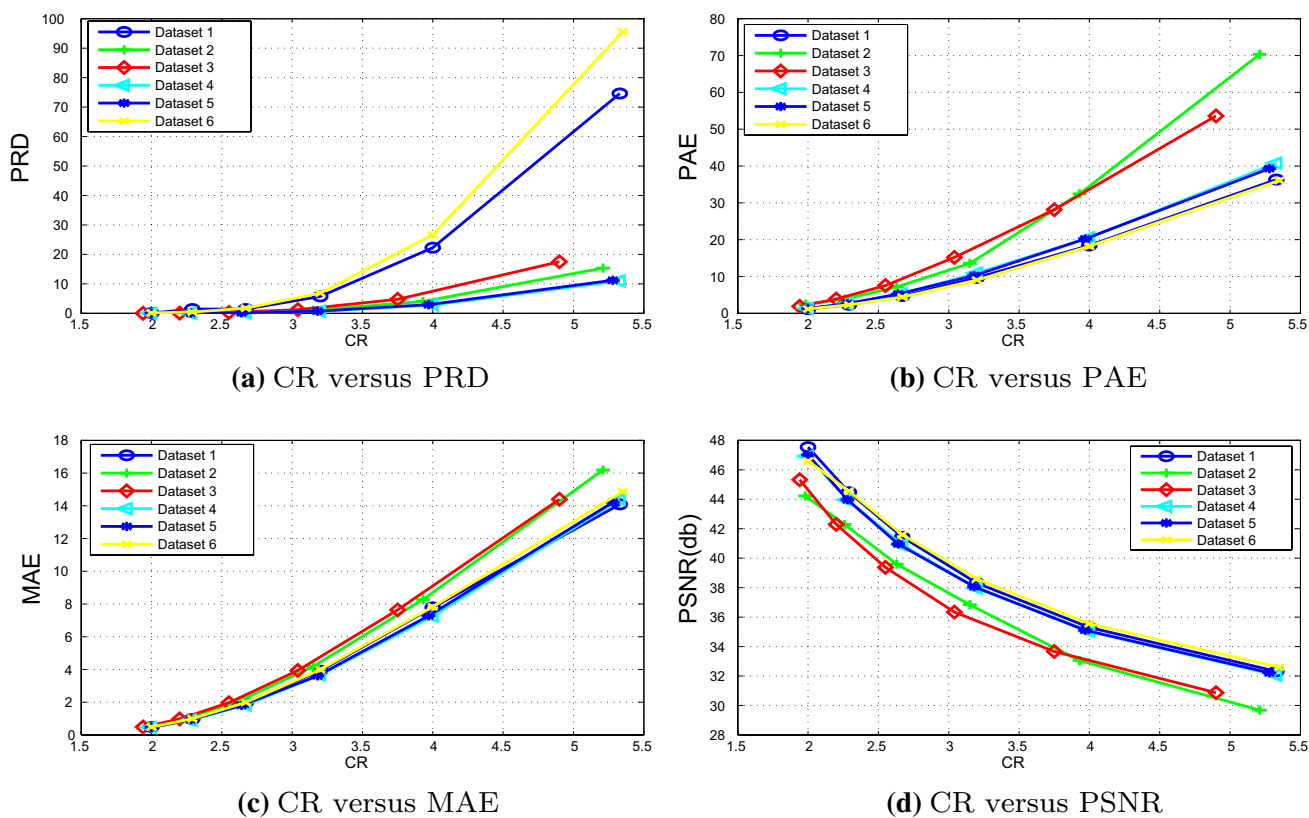


Fig. 6 Performance illustration of proposed scheme of different MCEEG data sets for varying CR

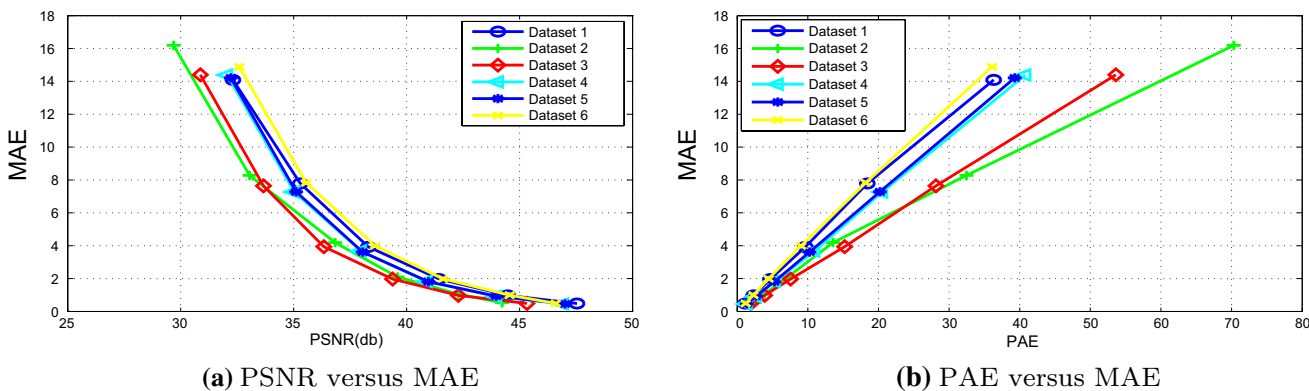


Fig. 7 Performance illustration of proposed scheme for different MCEEG data sets

from the table, in terms of compression and distortion parameters of the reconstructed signal. The performance of the proposed algorithm is compared with a very recent paper [36], which employs DPCM and clustering methodology to achieve compression and from the table it can be seen that the proposed algorithm outperforms both in CR and signal quality.

### Complexity analysis

In this section, the computational efficiency of the proposed scheme is analysed in terms of memory requirement and time complexity.

**Table 2** Computation time of the proposed scheme, for different sample sizes

No. of channels	No. of samples per channel			
	800		400	
	Encoding time (s)	Decoding time (s)	Encoding time (s)	Decoding time (s)
10	3.38	0.23	2.06	0.1
20	6.03	0.58	3.37	0.23
30	8.58	1.09	4.74	0.39
40	11.26	1.76	6.02	0.58
50	13.76	2.56	7.28	0.82
60	16.48	3.52	8.67	1.1
61	16.87	3.63	8.64	1.1

**Memory requirement analysis**

Consider  $n$  distinct 24 bit samples given by  $\binom{2^{24}}{n}$  and by information theoretical argument, the number of bits to represent the data is given by  $\log\binom{2^{24}}{n}$ , which can be approximated using the property in combination as  $\log\binom{2^{24}}{n} \geq (n * \log\frac{2^{24}}{n})$ , the size of  $n$  elements or samples. It can be proved that the proposed method can compress the data reasonably well. For  $n$  samples, the integer and fractional parts are processed separately.

*Integer part*

To take advantage of the redundancies introduced by the TL transform, RLE is performed where the encoder returns two values: one is the distinct integer and the second is its occurrence count. Consider that there are  $c$  distinct integers in the set of  $n$  integers, such that  $c$  satisfies the condition,

$1 \leq c \leq n$ . The number of bits required to represent  $c$  integers is given by

$$\ln\binom{2^8}{c} \geq c * \ln\left(\frac{2^8}{c}\right) \tag{10}$$

Memory requirement for occurrence or count of  $c$  integers is the same as (10). The overall memory requirement of RLE scheme is given by

$$2 * \ln\binom{2^8}{c} \geq 2 * c * \ln\left(\frac{2^8}{c}\right) \tag{11}$$

*Fraction part*

The data is binarized using  $d$  bits defined by bit depth. The number of samples  $n$  is taken such that product of  $d * n$  is divisible by eight (size of the integer). Following the similar procedures, the number of bits required for representing the fraction is given by  $2 * \log\binom{2^8}{k}$ , where  $k = \frac{d*n}{8}$ . To illustrate the effect of bit depth on memory requirements, consider 3 cases  $d$  takes values 16, 8, 1, the memory requirement is  $\log\binom{2^8}{2n}$ ,  $\log\binom{2^8}{n}$  and  $\log\binom{2^8}{\frac{n}{8}}$  respectively. It has been observed from the illustration that bit requirement reduces as the value of  $d$  decreases.

Total memory requirement of the proposed scheme is the sum of memory required for bit depth ( $d$ ), translation factor ( $f$ ), encoding of the integer and fractional part as shown in (12). The requirement of the scheme is much smaller than the actual requirement in storing single precision samples.

$$8 + 8 + 2 * c * \ln\left(\frac{2^8}{c}\right) + \frac{d * n}{8} \ln\left(\frac{2^8}{\frac{d * n}{8}}\right) \tag{12}$$

Variation in (12) occurs when

**Table 3** Relative performance comparison of the compression algorithm

	CR	PRD	PSNR	Optimal criteria
<i>Data set 4</i>				
Proposed algorithm	<b>5.33</b>	11.06	35.06	Bit depth of 3
Fast DCT [24]	4	11.09	****	BinDCT
<i>Data set 7</i>				
Proposed algorithm	<b>3.20</b>	4.43	38.66	Bit depth of 5
Wavelet image and volumetric coding [22]	2.56	1.72	37.95	Wavelet-s/s/t
Clustering method [36]	2.67	****	****	DPCM- KNN
<i>Data set 8</i>				
Proposed algorithm	3.61	8.73	19.40	Bit depth of 3
Wavelet image and volumetric coding [22]	<b>6.63</b>	9.21	28.92	Wavelet-s/s/t

\*\*\*\*The metric has not been evaluated in the algorithm



- Number of samples  $n$ , becomes very large leading to larger occurrence or count values;
- Translation factor  $f$  can take on larger values depending on the gain of MCEEG recorders.

An upper bound can be set to handle larger  $f$  and occurrence or count of distinct integers  $c$  in  $I_r(x)$ . Even with this upper bound, the increase in space requirement is not significant since

- The number of distinct integers  $c$  resulting from the TL transform is very small when compared to the total number of samples  $n$ .
- Increase in the number of bits for representing larger translation factor  $f$  does not create a significant change in overall requirement, as a larger sample space is considered.

### Time complexity

Time complexity is another quality measure signifying the computational performance of any algorithm. Most EEG compression algorithms available in the literature do not comment on their complexity measure, making a comparative study difficult. As most of these algorithms employ methodologies like PCA, ICA, Compressive Sensing or decomposition techniques like SVD, according to known literature, the lowest complexity attainable is  $O(n^2)$ .

The upper bound for the complexity of the proposed algorithm is given by  $O(nmd)$ , where  $n$  is the samples,  $m$  is the number of channels and  $d$  is the precision or bit depth. As  $m$  and  $d$  are fixed, they do not contribute to complexity. Thus, the algorithm has a linear dependence with the number of samples taken over a period, given by  $O(n)$ .

### Conclusion

The significance of this study is that it provides a computationally simple model for compressing MCEEG signals. In the proposed SPC system, the signal is normalized using the TL transform and the resulting coefficients are coded using IFC, ensuing in data compression. The proposed scheme achieved an average CR of 3.16 with a computational complexity of only  $O(n)$  and average encoding and decoding times per sample of 0.3 and 0.04 ms respectively. This is reasonably better than the other methods which rely on computationally intensive operations, to achieve similar CR. Moreover, the distortion level indicators such as PRD, MAE, PAE, RMSE and PSNR showed promising results to substantiate that the algorithm is suitable for MCEEG compression. As a future scope, the inter and intra-channel

redundancies can be exploited to increase CR further and adaptively choose optimal bit depth to attain higher CR across different data sets, without compromising on the existing complexity.

### Compliance with ethical standards

**Conflict of interest** The authors, have no conflict of interest, affiliations with or involvement in any organization or entity with any financial interest, or non-financial interest in the subject matter or materials discussed in this article.

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

### References

1. Nidermeyer E, Silva FLD (2005) Electroencephalography: basic principles, clinical applications and related fields, 5th edn. Lippincott Williams and Wilkins, Philadelphia
2. Jinu J, Titus G, Purushothaman S (2014) EEG-based automatic detection of drowsy state. In: Artificial intelligence and evolutionary algorithms in engineering systems, vol 324. Springer, New Delhi, pp 65–72
3. Vojkan Mihajlovic RV, Grundlehner B, Penders J (2015) Wearable, wireless EEG solutions in daily life applications: what are we missing? IEEE J Biomed Health Inform 19:569–571
4. Antoniol G, Tonella P (1997) EEG data compression techniques. IEEE Trans Biomed Eng 44:105–114
5. Sriraam N, Eswaran C (2006) Context based error modeling for lossless compression of EEG signals using neural networks. J Med Syst 30:439–448
6. Memon N, Kong X, Cinkler J (1999) Context-based lossless and near-lossless compression of EEG signals. IEEE Trans Biomed Eng 3:231–238
7. Sriraam N (2012) Correlation dimension based lossless compression of EEG signals. Biomed Signal Process Control 7:379–388
8. Wongsawat SY, Rao KR (2006) Integer sub-optimal Karhunen–Loeve transform for multi-channel lossless EEG compression. In: Proceedings of the IEEE 14th European Signal Processing Conference (EUSIPCO '06), Florence, Italy, Sept. 4–8
9. Akansu AN, Serdijn WA, Selesnick IW (2010) Emerging applications of wavelets: a review. Phys Commun 3:1–18
10. Srinivasan K, Dauwels J, Reddy MR (2011) A two-dimensional approach to lossless EEG compression. Biomed Signal Process Control 6:387–394
11. Cardenas Barrera JL, Cardenas Barrera JV, Rodriguez Valdivia E (2004) A wavelet-packet based algorithm for EEG signal compression. Med Inform Internet Med 29:15–27
12. Jolliffe I (ed) (2002) Principal component analysis, 2nd edn. Statistics. Springer, New York
13. Comon P (1994) Independent component analysis: a new concept? Signal Process 36:287–314
14. Zacharia A, Jai J, Titus G (2013) Automatic EEG artifact removal by independent component analysis using critical EEG rhythms. IEEE Int Conf Control Commun Comput 2013:364–367
15. Aviyente S (2007) Compressed sensing framework for EEG compression. In: Proceedings of IEEE Workshop on Statistical Signal Processing, Madison, WI, pp 181–184
16. Morabito FC et al (2013) Enhanced compressibility of EEG signal in Alzheimers disease patients. IEEE Sens J 13:3255–3262

17. Said A (2003) Arithmetic coding. *Lossless Compression Handbook*. Academic Press, New York, pp 447–455
18. Said A, Pearlman WA (1996) A new fast and efficient image codec based on set partitioning in hierarchical trees. *IEEE Trans Circuits Syst Video Technol* 6:243–250
19. Capurro I et al (2006) Low complexity, multichannel, lossless and near lossless EEG compression. In: *Proceedings of IEEE 22nd European Signal Processing Conference (EUSIPCO '14)*, Florence, Italy, Sept. 1–5, pp 2040–2044
20. Sriraam N (2009) Context based near lossless compression of EEG signals using neural network predictors. *Int J Electron Commun (AE)* 63:311–320
21. Dauwels J, Srinivasan K, Ramasubba Reddy M, Cichocki A (2013) Near lossless multichannel EEG compression based on matrix and tensor decompositions. *IEEE J Biomed Health Inform* 17:708–714
22. Srinivasan K, Dauwels J, Ramasubba Reddy M (2013) Multichannel EEG compression: wavelet-based image and volumetric coding approach. *IEEE J Biomed Health Inform* 17:113–120
23. Lei Lin JPC, Meng Y, Li ZB (2015) Multichannel EEG compression based on ICA and SPIHT. *Biomed Signal Process Control* 20:45–51
24. Darius Birvinskas IM, Jusas V, Damasevicius R (2015) Fast DCT algorithms for EEG data compression in embedded systems. *Comput Sci Inform Syst* 12:49–62
25. IEEE standard for floating-point arithmetic, IEEE Std. 754 (2008)
26. IEEE standard for floating-point arithmetic, IEEE Std. 754 (1985)
27. IEEE standard for radix independent floating point arithmetic, IEEE Std. 854 (1987)
28. Run-length colour encoding, ITU Std. Recommendation T.45 (02/00) (2000)
29. Lemire D, Boytsov L (2012) Decoding billions of integers per second through vectorization. [arXiv:1209.2137v6](https://arxiv.org/abs/1209.2137v6)
30. Tangermann M et al (2012) Review of the BCI competition IV. *Front Neurosci* 6:55
31. Benjamin Blankertz GC, Miller KR (2002) Classifying single trial EEG: towards brain computer interfacing. In: Becker S, Diettrich TG, Ghahramani Z (eds) *Advances in neural information processing systems 14 (NIPS 01)*. MIT Press, Cambridge
32. Blankertz B et al (2006) The BCI competition III: validating alternative approaches to actual BCI problems. *IEEE Trans Neural Syst Rehab Eng* 14:153–159
33. Naeem M et al (2006) Separability of four class motor imagery data using independent components analysis. *J Neural Eng* 3:208–216
34. Goldberger AL et al (2000) PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 101:E215–E220
35. Bazn Prieto C, Blanco Velasco M, Crdenas-Barrera J, Cruz Roldn F (2012) Analysis of tractable distortion metrics for EEG compression applications. *Physiol Meas* 33:1237–1247
36. Hejrati AFB, Mohammadi FA (2017) Efficient lossless multichannel EEG compression based on channel clustering. *Biomed Signal Process Control* 31:295–300