Proceedings on Automation in Medical Engineering
Proc AUTOMED, Vol 2, No 1 (2023): Article Ref 741
www.journals.infinite-science.de/automed/article/view/741

Tensor-based approximation of multichannel ECG sections

T. Schanze[1*]

[1] *IBMT, FB LSE, Technische Hochschule Mittelhessen (THM), Gießen, Germany*
[*] *Corresponding author, email: thomas.schanze@lse.thm.de*

Abstract: Tensors are multi-dimensional arrays which can be used to represent multi-dimensional data. Here we show how to approximate and compress multiple QRS-aligned multichannel ECG sections, each comprising P-wave, QRS-complex, and T-wave, by higher order singular value decomposition (HOSVD). We also introduce a method for the selection of components obtained by HOSVD for data approximation and a measure to quantify the quality of this tensor-based approximation of multi-dimensional data. The compression performance of this lossy data compression approach is assessed by the compression ratio.

I. Introduction

When data structures have more than two dimensions they can be represented by multi-dimensional arrays or tensors. The order of a tensor can be defined by the independent directions, i.e., the dimensionality of the data structure, required to describe it. Matrices are two-dimensional data structures that can be analyzed by singular value decomposition [1, 2]. Higher dimensional data structures or tensors with an order ≥ 3 can be decomposed in different ways. A well-known method is higher order singular value decomposition (HOSVD) [3, 4].

Akin to [5], we use HOSVD to approximate a multichannel biosignal, i.e., multiple QRS-aligned multichannel ECG sections, each comprising P-wave, QRS complex, and T-wave. We show how such a three-dimensional signal structure can be encoded by a third order tensor and introduce a simple method for selecting the elements of the so-called core tensor and the corresponding three matrices, all of which are obtained by HOSVD. In addition, we present a method to measure the quality of the tensor-based approximation of the data. Since HOSVD allows signal approximation, we use the compression ratio to quantify the performance of this lossy data compression approach.

II. Material and methods

II.I. Higher order singular value decomposition, data approximation and compression

Let $\mathbf{A}$ be a real $(m \times n)$ matrix of rank r. Its singular value decomposition (SVD) is

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathbf{T}}, \qquad (1)$$

where $\mathbf{U}$ is an orthogonal $(m \times m)$ matrix, $\mathbf{S}$ is a $(m \times n)$ diagonal matrix with non-negative entries, i.e., $s_{ii} \geq s_{(i+1)(i+1)} \geq 0$, and $\mathbf{V}$ is an orthogonal $(n \times n)$ matrix. Equation (1) can be written as follows:

$$\mathbf{A} = \sum_{i=1}^{r} s_{ii}\mathbf{u}_i\mathbf{v}_i^T = \sum_{i=1}^{r} s_{ii}\mathbf{u}_i \otimes \mathbf{v}_i, \qquad (2)$$

where $\mathbf{u}_i$ and $\mathbf{v}_i$ are the i-th columns of $\mathbf{U}$ and $\mathbf{V}$, $\otimes$ is the outer product. Note, $s_{ii} = 0$ for $i > r$, $r = \text{rank}\,(\mathbf{A})$. Furthermore, if $\mathbf{A}$ has full rank, then $r = \min(m, n)$. It is easy to verify that the larger an s_{ii} is, the more important is $s_{ii}\mathbf{u}_i \otimes \mathbf{v}_i$ for the representation of $\mathbf{A}$ by (2). A rank reduced approximation of order $q < r$ of $\mathbf{A}$ is

$$\mathbf{A} = \sum_{i=1}^{q} s_{ii}\mathbf{u}_i\mathbf{v}_i^T = \sum_{i=1}^{q} s_{ii}\mathbf{u}_i \otimes \mathbf{v}_i. \qquad (3)$$

Each $(m \times n \times p)$ tensor $\mathbf{T}$ of order 3 can be written in a similar way to the SVD of a matrix:

$$\mathbf{T} = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{o} s_{ijk}\mathbf{u}_i \otimes \mathbf{v}_j \otimes \mathbf{w}_k. \qquad (4)$$

Note, $\mathbf{U} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \dots \quad \mathbf{u}_m]$, $\mathbf{V} = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \dots \quad \mathbf{v}_n]$, and $\mathbf{W} = [\mathbf{w}_1 \quad \mathbf{w}_2 \quad \dots \quad \mathbf{u}_o]$ are unitary matrices, the s_{ijk} are the elements of the so called core tensor $\mathbf{S}$. Note, the product $\mathbf{u}_i \otimes \mathbf{v}_j \otimes \mathbf{w}_k$ has the same size as $\mathbf{T}$. An extension to orders > 3 is possible.

If $\mathbf{T}$ is a real, then the s_{ijk} are real, but not always positive. By analogy with the SVD, it is reasonable to conclude, that the larger $|s_{ijk}|$, the more important is $s_{ijk}\mathbf{u}_i \otimes \mathbf{v}_j \otimes \mathbf{w}_k$ for the representation of $\mathbf{T}$ by (4).

Let $s_{ijk;\,q} = s_{ijk}$ for the q largest $|s_{ijk}|$ and $s_{ijk;\,q} = 0$ for the remainder. Then

$$\mathbf{T}_q = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{o} s_{ijk;\,q}\mathbf{u}_i \otimes \mathbf{v}_j \otimes \mathbf{w}_k \qquad (5)$$

is an approximation of $\mathbf{T}$, which is a lossy compression.

Let $\#_{q;\,\mathbf{u}_i}$ be the number of $\mathbf{u}_i$ which are necessary to compute (5), then the number of values to calculate $\mathbf{T}_q$ is

$$\#_{\mathbf{T}_q} = \#_{q;\,\mathbf{u}_i}m + \#_{q;\,\mathbf{v}_i}n + \#_{q;\,\mathbf{w}_i}o + 4q. \qquad (6)$$

The term $4q$ is due to storage of $s_{ijk;\,q}$ and of the related indices. The compression ratio is

$$CR = m\,n\,o / \#_{\mathbf{T}_q}. \qquad (7)$$

A simple similarity measure of $\mathbf{T}$ and $\mathbf{T}_q$ is the Person correlation coefficient $\rho(\mathbf{T}, \mathbf{T}_q)$ of the vectorized $\mathbf{T}$ and $\mathbf{T}_q$.

II.II. Signals and programming

The 12-lead ECG data set A0170 of the China Physiological Signal Challenge 2018 was used [6]. Programming was done with Matlab (R022a) and a tensor toolbox [7]. The analyzed data set consists of 401 data points times 12 channels (Einthoven, Goldberger, Wilson) times 14 sections, i.e., 67,368 data points. The ECG data had been sampled @ 500 Hz. Sections were, for processing, QRS-aligned with respect to Einthoven lead 2. Each section comprises P-wave, QRS-complex and T-wave (Fig. 1).

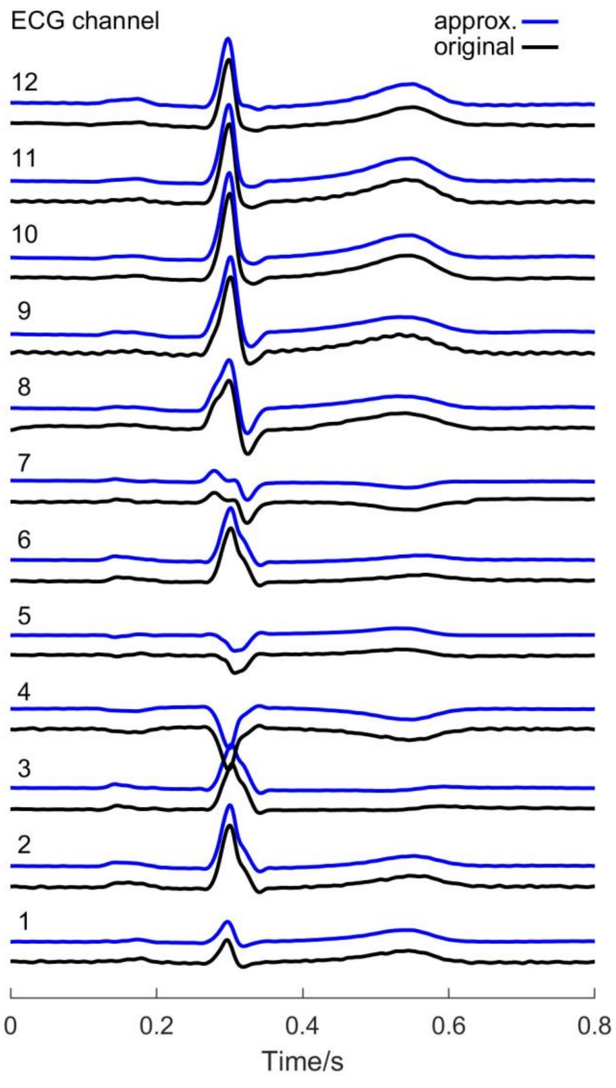

Figure 1: Example of a HOSVD-based data approximation. Shown is the fifth section of a multichannel ECG recording. Note, the high similarities between original and approximated signals. However, small hum-like ripples, such as those visible in the original channel 9, are not present in the approximation.

III. Results and discussion

Fig. 1 shows original and approximated time courses of the fifth section of the multichannel ECG recording used. The approximation was calculated for 10 elements of the core tensor, thus $q = 10$. These elements were selected, i.e., they are the core elements with the 10 largest absolute values. The sum of the absolute values of these 10 elements divided by the sum of the absolute values of all core tensor elements is 24.4 %. This combined with the similarities of original and approximated ECG indicates efficient coding.

The number of vectors of the matrices $\mathbf{U}, \mathbf{V},$ and $\mathbf{W}$ used for computing the approximation $\mathbf{T}_q$ of tensor $\mathbf{T}$, coding the original signals, are 7, 6, and 3, respectively. Thus, some vectors of $\mathbf{U}, \mathbf{V},$ and $\mathbf{W}$ were used multiple times to compute $\mathbf{T}_q$. This is another indication of the coding or compression potential of signals using HOSVD.

Using (6) we get for the number of data points necessary to compute the approximation: $\#_{\mathbf{T}_q} = 2{,}961$. Thus, the compression ratio is $CR = 67{,}368/2{,}961 \approx 22.8$.

For the Pearson correlation coefficient, calculated for the vectorized original data and the vectorized approximation, we get $\rho(\mathbf{T}, \mathbf{T}_q) = 0.9913$. The original data and the approximation are therefore very similar, as already indicated by Fig. 1.

The small hum-like ripples, clearly visible in channel 9 of the original data (Fig. 1), are not present in the approximation. However, this can be fixed by selecting more tensor core elements. For $q = 1{,}000$ the correlation increased to $\rho(\mathbf{T}, \mathbf{T}_q) = 0.9995$, but at the expense of compression ratio, which reduced to $CR = 67{,}368/19{,}129 \approx 3.5$.

The presented method for approximation and compression of ECG uses HOSVD. The selection of components to compute $\mathbf{T}_q$ is a successful extension of SVD-based methods [2]. The excellent approximation and compression properties of the method are surely due to the exceptional exploitation of the correlations present in the data (cf. [5]).

IV. Conclusions

HOSVD is a useful tool to approximate or compress ECG. The presented method promises high performance, but it should be compared with other approximation and compression methods and analyzed for other data in future work.

Acknowledgments

Fars Samann, thanks for discussions.

Author's statement

The author states: no funding involved and no conflict of interest.

References

[1] V. Klema, A. Laub, *The singular value decomposition: Its computation and some applications*, IEEE Transactions on Automatic Control, vol. 25, pp. 164–176, April 1980.

[2] T. Schanze, *Compression and Noise Reduction of Biomedical Signals by Singular Value Decomposition*, IFAC-PapersOnLine, vol. 51, pp. 361–366, 2018.

[3] T. G. Kolda, B. W. Bader, *Tensor Decompositions and Applications*, SIAM Rev., vol. 51, pp. 455–500, 2009.

[4] G. Bergqvist, E. G. Larsson, *The Higher-Order Singular Value Decomposition: Theory and an Application*, IEEE Signal Processing Magazine, vol. 27, pp. 151-154, May 2010.

[5] Y. S. Velchev, *An Approach for Multichannel ECG Compression Using HOSVD*, Proc. 28th Nat. Conf. with International Participation "Telecom 2020", October 29-30, 2020, Sofia, Bulgaria.

[6] F. Liu et al., *An open access database for evaluating the algorithms of ECG rhythm and morphology abnormal detection*, J. of Medical Imaging and Health Informatics, vol. 8, pp. 1368–1373, 2018.

[7] B. W. Bader et al., *Tensor Toolbox for MATLAB, Version 3.4*, www.tensortoolbox.org, September 21, 2022.