

Kent Academic Repository

Full text document (pdf)

Citation for published version

Jelfs, Beth and Li, Ling and Tin, Chung and Chan, Rosa H.M. (2016) Fuzzy entropy based nonnegative matrix factorization for muscle synergy extraction. In: UNSPECIFIED, 20-25 March 2016, Shanghai.

DOI

<https://doi.org/10.1109/ICASSP.2016.7471773>

Link to record in KAR

<http://kar.kent.ac.uk/55634/>

Document Version

Author's Accepted Manuscript

Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

Versions of research

The version in the Kent Academic Repository may differ from the final published version.

Users are advised to check <http://kar.kent.ac.uk> for the status of the paper. **Users should always cite the published version of record.**

Enquiries

For any further enquiries regarding the licence status of this document, please contact:

researchsupport@kent.ac.uk

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at <http://kar.kent.ac.uk/contact.html>

FUZZY ENTROPY BASED NONNEGATIVE MATRIX FACTORIZATION FOR MUSCLE SYNERGY EXTRACTION

Beth Jelfs^{1,2} Ling Li³ Chung Tin^{2,4} Rosa H.M. Chan^{1,2}

¹Dept. of Electronic Engineering, City University of Hong Kong, Hong Kong

²Centre for Biosystems, Neuroscience, & Nanotechnology, City University of Hong Kong, Hong Kong

³Dept. Computing, University of Kent, UK

⁴Dept. Mechanical & Biomedical Engineering, City University of Hong Kong, Hong Kong

E-mail: {beth.jelfs,rosachan,chungtin}@cityu.edu.hk, C.Li@kent.ac.uk

ABSTRACT

The concept of muscle synergies has proven to be an effective method for representing patterns of muscle activation. The number of degrees of freedom to be controlled are reduced while also providing a flexible platform for producing detailed movements using synergies as building blocks. It has previously been shown that small components of movement are crucial to producing precise and coordinated movement. Methods which focus on the variance of the data make it possible to overlook these small components in the synergy extraction process. However, algorithms which address the inherent complexity in the neuromuscular system are lacking. To that end we propose a new nonnegative matrix factorization algorithm which employs a cross fuzzy entropy similarity measure, thus, extracting muscle synergies which preserve the complexity of the recorded muscular data. The performance of the proposed algorithm is illustrated on representative EMG data.

Index Terms— Matrix Factorization, NMF, Fuzzy Entropy, Muscle Synergies, EMG

1. INTRODUCTION

The human body is capable of great dexterity and coordination, giving it the capability to produce a wide array of precise detailed movements but, as such, it is a highly complicated and also redundant system [1]. With such large numbers of degrees of freedom, how the central nervous system implements a control strategy is of interest in numerous different fields, from robotics to rehabilitation [2]. The concept of synergies, that is, common patterns of muscle activations which can be reused and serve as building blocks to produce detailed movements, has been gaining traction in recent years [3, 4, 5]. Synergies provide an approach which reduces the number of degrees of freedom to be controlled individually whilst providing the flexibility and adaptability to produce a wide range of motions.

Effective assessment of muscle synergies requires an accurate method for extracting the synergies. The standard approach to muscle synergy extraction is to use a nonnegative matrix factorization (NMF). That is, given a data matrix $Y \in \mathbb{R}^{X \times N}$ and positive integer $K < \min\{X, N\}$ find nonnegative matrices $W \in \mathbb{R}^{X \times K}$ and $H \in \mathbb{R}^{K \times N}$

$$\min_{W,H} D(Y||WH) \quad \text{subject to} \quad W \geq 0, \quad H \geq 0, \quad (1)$$

This work was supported by the Research Grants Council of the Hong Kong Special Administrative Region, China [Project No. CityU110813].

where $D(Y||WH)$ is an appropriately selected measure of goodness of fit [6], with the most common NMF algorithms being the alternating least squares (ALS) and the multiplicative Lee-Seung algorithms [7]. The consideration of the choice of algorithm used to extract muscle synergies has been limited to only a relatively small number of studies. For instance, the authors in [3] have compared different factorization algorithms including NMF, principal component analysis (PCA) and independent component analysis (ICA), and NMF algorithms based on signal dependent noise have been developed in [8] and [9].

While many of the approaches to identifying muscle synergies perform well in terms of producing a general factorization of the data, they do not always consider the full problem [10]. The success of the factorization is often based on assessing the variance accounted for (VAF) of the extracted synergies. However, success should be based not only on finding a representation which can, to a large extent, fit the data but it is also crucial to consider the task being performed. The reconstructed data should be capable of producing movement which could complete the required task [2, 11].

Of particular interest is the ability to produce not just movement in terms of the broad strokes but also the fine, precise movement which produces detailed coordinated actions. Our previous work has shown that improvements in coordination of finger movement are represented in the small components of the synergistic movement [12]. Therefore, if we want to successfully extract muscle activity which represents not only to the majority of the movement but also the subtle differences which account for better coordination we need to consider alternative measures as VAF may not be sufficient.

When it comes to the classification and assessment of changes in muscle activity recorded via surface electromyography (EMG), entropy is a commonly used tool [13, 14, 15], with measures based on approximate entropy (ApEn) [16] being a popular choice. ApEn provides a family of statistics to allow estimation of the rate of information generation in a time series, with sample entropy (SampEn) [17] being a modification of ApEn aimed at reducing the bias inherent in the calculation of ApEn. However, both ApEn and SampEn rely on an absolute decision boundary to determine whether two vectors are similar or not. In many systems, particularly biological systems which display a wide range of variability, such an absolute classification is not always realistic and the algorithm becomes sensitive to the choice of this decision boundary. In contrast fuzzy entropy (FuzzyEn) [18] employs a continuous fuzzy decision boundary to give a relative degree of similarity, resulting in a more consistent measure of entropy which is less sensitive to parameter selection [19]. Furthermore, FuzzyEn has been shown to be an ef-

fective measure of EMG data [18, 20].

As entropy measures have proven effective in characterizing EMG data, and any reconstructed data from a factorization of such EMG data must produce muscle activations capable of replicating the desired movement. It is therefore a natural extension to include an entropy measure into the muscle synergy extraction. Entropy rates have previously been used in blind source separation [21, 22], and for NMF corentropy [23] has been shown to provide an accurate similarity measure [24, 25, 26]. We therefore propose to include a FuzzyEn based measure into the update of the NMF algorithm.

The rest of this paper is organised as follows, in Section 2 we first introduce the cross FuzzyEn (C-FuzzyEn) [27] measure and then derive the update of a gradient descent NMF algorithm using a C-FuzzyEn similarity measure. In Section 3 we then test the algorithm for muscle synergy extraction on EMG data and compare the algorithm performance with that of the standard alternating least squares algorithm. Finally, Section 4 concludes the paper.

2. FUZZY ENTROPY NMF

One of the keys to successful NMF is the selection of the distance measure between the data matrix, Y , and the decomposed matrices, W and H . Standard measures such as the l_2 norm distance or the Kullback-Leibler divergence are linear similarity measures and may not be appropriate in situations where the data contains nonlinearities or are extremely noisy. With that in mind, and in order to preserve the complexity of the data, our algorithm employs the C-FuzzyEn between Y and $W \cdot H$ as a similarity measure between the data and the decomposition. The measure is first applied row-wise between Y and $W \cdot H$ in order to update W , followed by column-wise for the update of H . Such an approach ensures both the individual muscle activity and the interaction between the muscles is represented as accurately as possible by the decomposition.

2.1. Cross Fuzzy Entropy

FuzzyEn was introduced to overcome limitations of ApEn and SampEn due to their reliance on the use of the Heaviside function. As discussed previously both algorithms employ an absolute boundary in their determination of whether two vectors are similar or not. In contrast FuzzyEn employs an exponential function to give a continuous degree of similarity between vectors based on their closeness [18]. As a natural extension of FuzzyEn, C-FuzzyEn has been introduced in order to assess the degree of similarity between two different signals using fuzzy entropy [27, 28]. The C-FuzzyEn of two time series of N points,

$$\{u(i) : 1 \leq i \leq N\} \quad (2)$$

$$\{v(j) : 1 \leq j \leq N\}, \quad (3)$$

is estimated from the $N - m$ vectors of length m ,

$$\mathbf{x}_m(i) = \{u(i+k) : 0 \leq k \leq m-1\} - \bar{u}(i) \quad (4)$$

$$\mathbf{y}_m(j) = \{v(j+k) : 0 \leq k \leq m-1\} - \bar{v}(j), \quad (5)$$

where the mean values of the vectors, $\bar{u}(i)$ and $\bar{v}(j)$, are removed to eliminate any baseline and allow comparison of their similarity based only on their shape rather than absolute position. The distances between any pair of vectors is given in terms of the infinity norm, that is the maximum absolute difference between their scalar components

$$d_{i,j}^m = \|\mathbf{x}_m(i) - \mathbf{y}_m(j)\|_\infty. \quad (6)$$

Their degree of similarity can then be defined in terms of a fuzzy function of this distance. In this case the exponential function $\exp -(\cdot)^p / r$, where p defines the gradient of the boundary and r the width of the boundary, such that

$$D_{i,j}^m(p, r) = \exp\left(-\frac{(d_{i,j}^m)^p}{r}\right). \quad (7)$$

The overall similarity function for vectors of length m , ϕ^m , then becomes

$$\phi^m(p, r) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left(\frac{1}{N-m} \sum_{j=1}^{N-m} D_{ij}^m \right), \quad (8)$$

and in the same vein the similarity function of vectors of length $m+1$ is given by

$$\phi^{m+1}(p, r) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left(\frac{1}{N-m} \sum_{j=1}^{N-m} D_{ij}^{m+1} \right). \quad (9)$$

The C-FuzzyEn of the two time series can be expressed in terms of the ratio of the negative log of the conditional probability ϕ^{m+1} / ϕ^m which for a finite time series is equivalent to

$$\text{C-FuzzyEn}(m, p, r, N) = \ln \phi^m(p, r) - \ln \phi^{m+1}(p, r). \quad (10)$$

For simplicity the terms C-FuzzyEn(m, p, r, N), $\phi^m(p, r)$ and $\phi^{m+1}(p, r)$ will be referred to as C-FuzzyEn, ϕ^m and ϕ^{m+1} in the following derivation.

2.2. Proposed NMF Update

First we define the measure of goodness of fit $D(Y||WH)$ in the NMF algorithm (1) in terms of the C-FuzzyEn (10) of the elements of Y and $W \cdot H$. Then using a gradient descent rule, where the gradient of the cost function is taken row-wise with respect to the individual elements of W , and column-wise with respect to the elements of H , we have the updates

$$[W]_{x,k} \leftarrow W_{x,k} - \eta^W \nabla_{[W]_{x,k}} \left(D(Y||WH) \right), \quad (11)$$

$$[H]_{k,n} \leftarrow H_{k,n} - \eta^H \nabla_{[H]_{k,n}} \left(D(Y||WH) \right), \quad (12)$$

where η^W and η^H are the respective learning rates. For simplicity, we will consider the cost function in terms of its constituent parts, $\ln \phi^m$ and $\ln \phi^{m+1}$, and take the gradient of each separately.

For the row-wise measure of C-FuzzyEn we substitute

$$\mathbf{y}_x^m(i) = \{[Y]_{x,i}, \dots, [Y]_{x,i+m-1}\} - [\bar{Y}]_{x,i}^m, \quad (13)$$

$$\mathbf{wh}_x^m(j) = \{[WH]_{x,j}, \dots, [WH]_{x,j+m-1}\} - [\overline{WH}]_{x,j}^m, \quad (14)$$

into the distance measure (6). Then taking the gradient of $\ln \phi_x^m$ with respect to the components of W gives

$$\begin{aligned} \nabla_{[W]_{x,k}} \ln \phi_x^m &= \frac{1}{\phi_x^m} \sum_{i=1}^{N-m} \sum_{j=1}^{N-m} \left(\frac{p}{r} D_{i,j}^m \left(d_{i,j}^m \right)^{p-1} \right. \\ &\quad \cdot \left. \left([H]_{k,j+c^m} - [\bar{H}]_{k,j}^m \right) \text{sgn} \left(\mathbf{y}_x^m(i)(c^m) - \mathbf{wh}_x^m(i)(c^m) \right) \right), \end{aligned} \quad (15)$$

where $c^m \in [0, m-1]$ such that the c^{mth} elements of $\mathbf{y}_x^m(i)$ and $\mathbf{wh}_x^m(j)$, that is $\mathbf{y}_x^m(i)(c^m)$ and $\mathbf{wh}_x^m(j)(c^m)$ respectively, are the

elements which correspond to the maximum absolute difference between the vectors

$$\left\| \mathbf{y}_x^m(i)(c^m) - \mathbf{wh}_x^m(j)(c^m) \right\| = \left\| \mathbf{y}_x^m(i) - \mathbf{wh}_x^m(j) \right\|_{\infty}. \quad (16)$$

In the same manner if we consider the column-wise measure of C-FuzzyEn, calculating the distance measure between the vectors

$$\mathbf{y}_n^m(i) = \{[Y]_{i,n}, \dots, [Y]_{i+m-1,n}\} - [\overline{Y}]_{i,n}^m, \quad (17)$$

and

$$\mathbf{wh}_n^m(j) = \{[WH]_{j,n}, \dots, [WH]_{j+m-1,n}\} - [\overline{WH}]_{j,n}^m, \quad (18)$$

taking the gradient of $\ln \phi^m$ with respect to the components of H results in

$$\begin{aligned} \nabla_{[H]_{k,n}} \ln \phi_n^m &= \frac{1}{\phi_n^m} \sum_{i=1}^{X-m} \sum_{j=1}^{X-m} \left(\frac{p}{r} D_{i,j}^m (d_{i,j}^m) \right)^{p-1} \\ &\cdot \left([W]_{j+c^m,k} - [\overline{W}]_{j,k}^m \right) \operatorname{sgn} \left(\mathbf{y}_x^m(i)(c^m) - \mathbf{wh}_x^m(i)(c^m) \right). \end{aligned} \quad (19)$$

Following the same approach to calculate the gradients with respect to both W and H for $\ln \phi^{m+1}$ gives the gradient of the C-FuzzyEn cost function with respect to the components of W as

$$\begin{aligned} \nabla_{[W]_{x,k}} \ln \phi_x^m &= \nabla_{W_{x,k}} \ln \phi_x^m - \nabla_{W_{x,k}} \ln \phi_x^{m+1} \\ &= \frac{1}{\phi_x^m} \sum_{i=1}^{N-m} \sum_{j=1}^{N-m} \left(\frac{p}{r} D_{i,j}^m (d_{i,j}^m) \right)^{p-1} \left([H]_{k,j+c^m} - [\overline{H}]_{k,j}^m \right) \\ &\cdot \operatorname{sgn} \left(\mathbf{y}_x^m(i)(c^m) - \mathbf{wh}_x^m(i)(c^m) \right) - \frac{1}{\phi_x^{m+1}} \\ &\sum_{i=1}^{N-m} \sum_{j=1}^{N-m} \left(\frac{p}{r} D_{i,j}^{m+1} (d_{i,j}^{m+1}) \right)^{p-1} \left([H]_{k,j+c^{m+1}} - [\overline{H}]_{k,j}^{m+1} \right) \\ &\cdot \operatorname{sgn} \left(\mathbf{y}_x^{m+1}(i)(c^{m+1}) - \mathbf{wh}_x^{m+1}(i)(c^{m+1}) \right), \end{aligned} \quad (20)$$

and with respect to the components of H as

$$\begin{aligned} \nabla_{[H]_{k,n}} \ln \phi_n^m &= \nabla_{H_{k,n}} \ln \phi_n^m - \nabla_{W_{x,k}} \ln \phi_n^{m+1} \\ &= \frac{1}{\phi_n^m} \sum_{i=1}^{X-m} \sum_{j=1}^{X-m} \left(\frac{p}{r} D_{i,j}^m (d_{i,j}^m) \right)^{p-1} \left([W]_{j+c^m,k} - [\overline{W}]_{j,k}^m \right) \\ &\cdot \operatorname{sgn} \left(\mathbf{y}_n^m(i)(c^m) - \mathbf{wh}_n^m(i)(c^m) \right) - \frac{1}{\phi_n^{m+1}} \\ &\sum_{i=1}^{X-m} \sum_{j=1}^{X-m} \left(\frac{p}{r} D_{i,j}^{m+1} (d_{i,j}^{m+1}) \right)^{p-1} \left([W]_{j+c^{m+1},n} - [\overline{W}]_{j,k}^{m+1} \right) \\ &\cdot \operatorname{sgn} \left(\mathbf{y}_n^{m+1}(i)(c^{m+1}) - \mathbf{wh}_n^{m+1}(i)(c^{m+1}) \right). \end{aligned} \quad (21)$$

Substituting (20) and (21) into (11) and (12) respectively gives us the NMF update with a C-FuzzyEn based cost function distance measure.

Table 1. Muscles recorded on each channel and location over muscle for muscles with multiple recordings - upper indicating closer to the body (origin) and lower closer to the hand (insertion).

Channel	Muscle	Location
1	Extensor digitorum	upper
2	Anconeus	
3	Flexor carpi ulnaris	
4	Pronator teres	lower
5	Flexor carpi radialis	upper
6	Flexor carpi radialis	lower
7	Palmaris longus	
8	Pronator teres	upper
9	Extensor carpi ulnaris	
10	Extensor digitorum	lower
11	Extensor carpi radialis brevis	
12	Extensor carpi radialis longus	
13	Abductor pollicis brevis	
14	Abductor digiti minimi	
15	Biceps brachii	upper
16	Biceps brachii	lower

3. ALGORITHM PERFORMANCE

3.1. EMG Data

The EMG data has previously been described in [29, 30], in brief the data was recorded from 5 healthy right-handed subjects, all of whom gave informed consent prior to the experiment. The EMG data was acquired using a g.USBamp (Guger Technologies, Graz, Austria) amplifier. The data was sampled at 4.8kHz and before analysis was bandpass filtered between 10-200Hz, rectified and down-sampled by a factor of 10. Surface electrodes were placed over 12 different muscles located in the upper arm, forearm and hand of the subjects, with 4 of the larger muscles having electrodes placed over 2 different locations giving a total of 16 channels (listed in Table 1) with a reference electrode placed on the subjects earlobe. The experimental procedure involved the subject performing a series of simple hand and wrist movements. For the purposes of illustrating the performance of our algorithm the trials representing movement related to grasping a cylindrical object were selected. For each subject the grasping movement was repeated 3 times which were manually divided into separate trials giving a total of 15 trials analysed.

3.2. Algorithm Performance

The C-FuzzyEn has three parameters which need to be set. The parameter, m , is the length of the vectors to be compared, a larger value of m may result in a more detailed representation of the dynamics but also requires either large quantities of data - which can be difficult to obtain for biological datasets, or a wider decision boundary which will consequently lead to a loss in information. An appropriate choice of vector length for short biological datasets is a value of $m = 2$. The parameters p and r relate to the gradient and width, respectively, of the exponential function determining the fuzzy boundary. Too large and we risk information loss, and too small and the influence of noise is increased. Typically for p small integer values greater than 1 are used [19], in this case $p = 2$. For experimental data it can be beneficial to set the value of r as a function of the standard deviation of the data, thus, allowing comparison of data with different amplitudes [17]. In this case as we consider the updates of W and H in terms of the rows and columns of Y , for the updates

Table 2. Comparison of the average AIC for the proposed FuzzyEn NMF and the ALS NMF across a range of number of synergies.

Algorithm $\times 10^3$	Number of Synergies												
	3	4	5	6	7	8	9	10	11	12	13	14	15
ALS	9.397	12.528	15.659	18.790	21.992	25.053	28.184	31.316	34.448	37.579	40.711	43.848	46.977
FuzzyEn	9.390	12.521	15.652	18.783	21.915	25.045	28.177	31.308	34.439	37.570	40.702	43.831	46.963

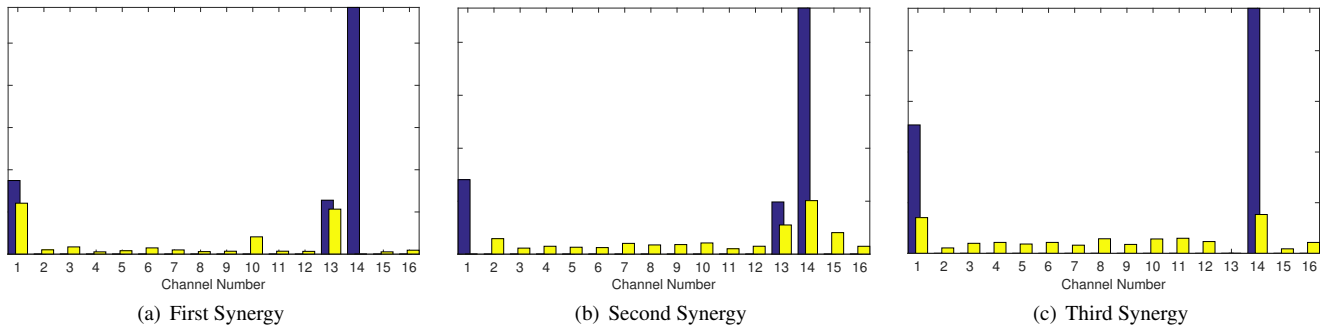


Fig. 1. Sample extracted synergies for $k=3$, blue bars - FuzzyEn NMF, yellow bars - ALS NMF.

of W , $r = 0.2 * \sigma_{Y_x}$ and for the updates of H , $r = 0.2 * \sigma_{Y_n}$. The learning rates of the NMF were determined empirically with $\eta^W = 0.8$ and $\eta^H = 0.1$.

The performance of the proposed algorithm was compared against a standard alternating least squares (ALS) NMF algorithm. To compare how well the decomposition from each algorithm fit the data, the Akaike Information Criteria (AIC) [31] was used

$$AIC(\theta) = -2L + 2n_p, \quad (22)$$

where L is the maximum log-likelihood and n_p is the number of independent parameters in the model. As the true number of synergies is unknown we compared both algorithms across a range of different numbers of bases, k . The AIC was determined for the lowest residual error between the extraction and the original data for each value of k for both algorithms. The average of the AIC across all trials for both algorithms, with the number of synergies ranging from 3 to 15, are shown in Table 2. These results indicate that, across all synergies, the proposed algorithm outperforms the standard ALS algorithm giving a better representation of the data in terms of the AIC.

In relation to muscle synergy extraction, the basis vectors contained within each column of the matrix W represent the synergy coefficients illustrating how much each muscle contributes to the synergy. The matrix H represents the time-varying coefficients of each basis across all the samples. To further compare the performance of the two algorithms we consider differences/similarities in the synergy coefficients contained within the W matrices extracted via each algorithm. Figure 1 provides a representative illustration of the synergies extracted for $k = 3$ for a single trial. Importantly, although there was only a small improvement in the AIC of the proposed algorithm over the ALS, as can be seen, there was a noticeable difference in the synergies extracted. The two different algorithms provide a significantly different representation of how the muscles contributed to the synergies. With the ALS NMF having a much more even spread of the contributions across all of the muscles. Whereas, the FuzzyEn NMF considered three muscles in particular

- extensor digitorum (upper), abductor pollicis brevis and abductor digiti minimi - to have larger fixed contributions to the synergies. These muscles correspond to the extension of the 4 long fingers, abduction/extension of the thumb and abduction/extension of the little finger, all muscles which would contribute to a grasping action. This highlights that there are significant differences between the algorithms despite the relatively modest increase in performance and that the FuzzyEn NMF may provide a significant decomposition for muscle synergy extraction.

4. CONCLUSIONS

We have presented a new NMF algorithm which incorporates C-FuzzyEn as the similarity measure in the update of the algorithm. The use of C-FuzzyEn allows us to decompose matrices whilst also preserving the complexity of our data. The preservation of this underlying dynamic structure is important when considering data for which standard measures of the variability within the data are insufficient. In particular we believe this to be the case for EMG data where small variability in the data contributes significantly to the quality and coordination of movements. We have highlighted the performance of the algorithm in comparison to a standard NMF algorithm for a representative EMG data set. The results illustrate that even small improvements in performance can significantly change the representation of the data. While we have shown the applicability of this algorithm for the factorization of EMG data sets, making it ideally suited to the extraction of muscle synergies, we also believe the algorithm could find application in other areas where the data under consideration contains small but significant fluctuations.

5. REFERENCES

- [1] N.A. Bernstein, *The Coordination and Regulation of Movements*, Pergamon Press, Oxford, UK, 1967.
- [2] C. Alessandro, I. Delis, F. Nori, S. Panzeri, and B. Berret,

- “Muscle synergies in neuroscience and robotics: from input-space to task-space perspectives,” *Front. Comput. Neurosci.*, vol. 7, pp. 43, 2013.
- [3] M.C. Tresch, V.C.K. Cheung, and d’Avella A., “Matrix factorization algorithms for the identification of muscle synergies: Evaluation on simulated and experimental data sets,” *J. Neurophysiol.*, vol. 95, no. 4, pp. 2199–2212, 2005.
- [4] A. d’Avella and E. Bizzi, “Shared and specific muscle synergies in natural motor behaviors,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 8, pp. 3076–3081, 2005.
- [5] G. Torres-Oviedo and L.H. Ting, “Subject-specific muscle synergies in human balance control are consistent across different biomechanical contexts,” *J. Neurophysiol.*, vol. 103, no. 6, pp. 3084–3098, 2010.
- [6] Andrzej Cichocki, Rafal Zdunek, Anh Huy Phan, and Shun-ichi Amari, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*, Wiley-Blackwell, Chichester, UK, 2009.
- [7] D.D. Lee and H.S. Seung, “Learning the parts of objects by non-negative matrix factorization,” *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.
- [8] V.C.K. Cheung and M.C. Tresch, “Non-negative matrix factorization algorithms modeling noise distributions within the exponential family,” in *Proc. Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society*, Jan 2005, pp. 4990–4993.
- [9] K. Devarajan and Vincent C.K. Cheung, “On nonnegative matrix factorization algorithms for signal-dependent noise with application to electromyography data,” *Neural Comput.*, vol. 26, no. 6, pp. 1128–1168, 2014.
- [10] M.C. Tresch and A. Jarc, “The case for and against muscle synergies,” *Curr. Opin. Neurobiol.*, vol. 19, no. 6, pp. 601–607, 2009.
- [11] I. Delis, B. Berret, T. Pozzo, and S. Panzeri, “Quantitative evaluation of muscle synergy models: a single-trial task decoding approach,” *Front. Comput. Neurosci.*, vol. 7, pp. 8, 2013.
- [12] B. Jelfs, S. Zhou, B.K.Y. Wong, C. Tin, and Chan R.H.M., “Recruitment of small synergistic movement makes a good pianist,” in *Proc. Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society*, Aug. 2015, pp. 242–245.
- [13] M. Kaufman, Ulrich Zurcher, and Paul S. Sung, “Entropy of electromyography time series,” *Physica A*, vol. 386, no. 2, pp. 698–707, 2007.
- [14] D.E. Vaillancourt and K.M. Newell, “The dynamics of resting and postural tremor in parkinson’s disease,” *Clin. Neurophysiol.*, vol. 111, no. 11, pp. 2046–2056, 2000.
- [15] P. Zhou and X. Zhang, “A novel technique for muscle onset detection using surface EMG signals without removal of ECG artifacts,” *Physiol. Meas.*, vol. 35, no. 1, pp. 45–54, 2013.
- [16] S.M. Pincus, “Approximate entropy as a measure of system complexity,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 88, no. 6, pp. 2297–2301, 1991.
- [17] J.S. Richman and J.R. Moorman, “Physiological time-series analysis using approximate entropy and sample entropy,” *Am. J. Physiol.: Heart Circ. Physiol.*, vol. 278, no. 6, pp. H2039–H2049, 2000.
- [18] W. Chen, Z. Wang, H. Xie, and W. Yu, “Characterization of surface EMG signal based on fuzzy entropy,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 15, no. 2, pp. 266–272, 2007.
- [19] W. Chen, J. Zhuang, W. Yu, and Z. Wang, “Measuring complexity using fuzzyen, apen, and sampen,” *Med. Eng. Phys.*, vol. 31, no. 1, pp. 61–68, 2009.
- [20] M. Lyu, C. Xiong, Q. Zhang, and L. He, “Fuzzy entropy-based muscle onset detection using electromyography (EMG),” in *Intelligent Robotics and Applications*, X. Zhang, H. Liu, Z. Chen, and N. Wang, Eds., vol. 8917 of *Lecture Notes in Computer Science*, pp. 89–98. Springer Science + Business Media, 2014.
- [21] G.-S. Fu, R. Phlypo, M. Anderson, X.-L. Li, and T. Adali, “Blind source separation by entropy rate minimization,” *IEEE Trans. Signal Process.*, vol. 62, no. 16, pp. 4245–4255, 2014.
- [22] G.-S. Fu, R. Phlypo, M. Anderson, and T. Adali, “Complex independent component analysis using three types of diversity: Non-Gaussianity, nonwhiteness, and noncircularity,” *IEEE Trans. Signal Process.*, vol. 63, no. 3, pp. 794–805, 2015.
- [23] I. Santamaria, P.P. Pokharel, and J.C. Principe, “Generalized correlation function: definition, properties, and application to blind equalization,” *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2187–2197, 2006.
- [24] T. Ensari, J. Chorowski, and J.M. Zurada, “Correntropy-based document clustering via nonnegative matrix factorization,” in *Artificial Neural Networks and Machine Learning ICANN 2012*, Alessandro E. P. Villa Villa, Włodzisław Duch, Péter Érdi, Francesco Masulli, and Günther Palm, Eds., vol. 7553 of *Lecture Notes in Computer Science*, pp. 347–354. Springer Science + Business Media, 2012.
- [25] J.J.-Y. Wang, X. Wang, and X. Gao, “Non-negative matrix factorization by maximizing correntropy for cancer clustering,” *BMC Bioinf.*, vol. 14, no. 1, pp. 107, 2013.
- [26] E.H. Asl and J.M. Zurada, “Multiplicative algorithm for correntropy-based nonnegative matrix factorization,” *J. Appl. Comput. Sci. Methods*, vol. 5, no. 2, pp. 89–104, 2013.
- [27] H.-B. Xie, Y.-P. Zheng, and J.-Y. Guo, “Detection of synchrony in biosignals using cross fuzzy entropy,” in *Proc. Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society*, Sept. 2009, pp. 2971–2974.
- [28] H.-B. Xie, Y.-P. Zheng, J.-Y. Guo, and X. Chen, “Cross-fuzzy entropy: A new method to test pattern synchrony of bivariate time series,” *Inform. Sciences*, vol. 180, no. 9, pp. 1715–1724, 2010.
- [29] L. Li, *Heterogeneous data fusion for brain psychology applications*, Ph.D. thesis, Imperial College London, 2011.
- [30] L. Li, D. Looney, C. Park, N.U. Rehman, and D.P. Mandic, “Power independent EMG based gesture recognition for robotics,” in *Proc. Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society*, Aug. 2011, pp. 793–796.
- [31] H. Akaike, “A new look at the statistical model identification,” *IEEE Trans. Automat. Contr.*, vol. 19, no. 6, pp. 716–723, 1974.