

International Journal of Neural Systems, Vol. 0, No. 0 (0) 1–23
© World Scientific Publishing Company

FEATURE EXTRACTION WITH GMDH-TYPE NEURAL NETWORKS FOR EEG-BASED PERSON IDENTIFICATION

Vitaly Schetinin^{1*}, Livija Jakaite¹, Ndifreke Nyah¹, Dusica Novakovic¹, Wojtek Krzanowski²

¹*School of Computer Science and Technology, University of Bedfordshire, Park Square,
Luton, UK*

vitaly.schetinin@beds.ac.uk

²*College of Engineering, Mathematics and Physical Sciences, University of Exeter,
Exeter, UK*

The brain activity observed on EEG electrodes is influenced by volume conduction and functional connectivity of a person performing a task. When the task is a biometric test the EEG signals represent the unique “brain print”, which is defined by the functional connectivity that is represented by the interactions between electrodes, whilst the conduction components cause trivial correlations. Orthogonalisation using autoregressive modelling minimises the conduction components, and then the residuals are related to features correlated with the functional connectivity. However the orthogonalisation can be unreliable for high-dimensional EEG data. We have found that the dimensionality can be significantly reduced if the baselines required for estimating the residuals can be modelled by using relevant electrodes. In our approach, the required models are learnt by a Group Method of Data Handling (GMDH) algorithm which we have made capable of discovering reliable models from multidimensional EEG data. In our experiments on the EEG-MMI benchmark data which include 109 participants, the proposed method has correctly identified all the subjects and provided a statistically significant ($p < 0.01$) improvement of the identification accuracy. The experiments have shown that the proposed GMDH method can learn new features from multi-electrode EEG data, which are capable to improve the accuracy of biometric identification.

Keywords: Biometrics, multi-electrode EEG, Brain functional connectivity, Volume conduction, Feature extraction, Group Method of Data Handling

1. Introduction

Recent advances in neural engineering and human-machine interaction based on the electroencephalogram (EEG) are receiving much attention and active development in many application areas. In biometric security applications, EEG signals cannot be reproduced by an intruder or remotely captured with sensors^{1,2} and are extremely difficult to imitate. A person cannot be forced to reproduce a biometric test under stress conditions.³ The recently developed EEG sensor technologies have significantly improved

the usability of EEG headsets, thus making EEG-based technologies user-friendly.⁴

EEG-based person identification and recognition methods employ different approaches to extracting EEG features that can represent a person’s individuality as a “brain print”. The EEG features are typically represented by frequency spectra.^{5–7} When EEG recordings are made from a multi-electrode system, the features are extracted for each electrode.

New promising approaches^{8–10} to EEG biometrics are based on the “connectome” which reflects individual differences in the brain organisation, known

*Corresponding author.

as the brain functional connectivity, of a person.^{11,12}

The connectivity represents the interaction between different brain regions.¹³ Attempts have been made to estimate the interactions in terms of correlations between EEG electrodes.^{14,15} However, the correlations can be misleading for estimating the connectivity because of the volume conduction problem. The volume conduction causes the neural activity from a single source to be received by multiple electrodes. As the result, the observed correlations between the electrodes are trivial and do not reflect the true connectivity.

Following,¹⁶ an event can be caused by volume conduction if (i) the event is seen at more than two electrodes, (ii) the event appears with the same waveform, and (iii) peaks and deflections occur synchronously at all electrodes where the event is observed.

Recently, it has been proposed to estimate the connectivity by computing the phase shifts between EEG signals.^{9,10} It has been shown that phase shifts cannot be consistently caused by the volume conduction, and so are predominately correlated with the true connectivity.^{17,18}

However, the phase differences are computed for each pair of electrodes, and the number of pairs can be large for multi-electrode systems. The delays are estimated for each frequency that can be present in the EEG signal. This makes the brain print data high-dimensional.

Orthogonalisation using Autoregressive (AR) modelling minimises the trivial correlations, and the connectivity features can be then extracted from the residuals.^{8,19,20} However this approach can be unreliable in the case of high-dimensional EEG data.

The desired solutions can be found with conventional machine learning methods in cases where model structures are well-defined. These methods are limited in terms of modelling real-world problems represented by underdetermined, “ill-defined” or contaminated data, see e.g.²¹ Under such conditions conventional Artificial Neural Networks (ANN) are prone to providing unreliable solutions.^{22–24}

The other machine learning methods such as regression and classification trees,²⁵ which are based on recursive data partitioning, are limited in their ability to generalise, their prediction accuracy can be poor on new data.

The above problems with model structures and

data have been partly resolved within the Group Method of Data Handling (GMDH),^{22,26,27} shown to deliver efficient data-driven solutions to problems of modelling, prediction, and pattern recognition. Nowadays GMDH algorithms are considered as Deep Learning paradigms due to their ability to generate new features for multi-layered ANN, which makes their architectures “deep”.²⁸ Within the Deep Learning framework, neurons at new layers of neural networks are generated with new features which are capable of improving the ability to generalise.^{28–30}

The advantages of the GMDH approach have been shown in our early research,^{31,32} in which a new algorithm developed for learning GMDH-type ANN has outperformed the conventional ANN trained with the back-propagation method. The methods have been compared in terms of the accuracy of classification of clinical EEG.

In this paper we aim (i) to explore the conditions under which reliable biometric features can be extracted from the residuals, which are associated with the brain functional connectivity individual for a person, and (ii) to develop a new feature extraction method for EEG-based person identification using the advantages of Deep Learning.

In sections 2 and 3 we review the related literature, define the problem and consider limitations of the existing methods. In two other sections 4 and 5 we describe the proposed method and EEG data for our experiments. In sections 6 and 7 we present the experiments and discuss the results, and finally section 8 concludes the paper.

2. Related Work

EEG-based person identification and recognition methods employ different approaches to extracting EEG features that can represent a person’s individuality as a “brain print”. EEG features are represented by frequency spectra, which are calculated by using Fourier or wavelet transforms, or by the coefficients of AR models. EEG records are typically made from a multi-electrode EEG system during the specified types of activities performed by the subject.

The most recent EEG-based biometric studies have been aimed at extracting the “connectome” which represents the brain functional connectivity of a person. EEG signals, which are recorded from electrodes in different positions on the scalp, contain the important features that can be extracted for

purposes of estimating the connectivity and person identification.

In this section we review the existing methods of EEG-based identification, which have been validated on representative data sets. We attempt to analyse the main approaches to estimating the brain functional connectivity because we think that they are the most promising.

2.1. Spectral-based features

One of the earliest attempts to develop an EEG-based person recognition system was based on the statistical analysis of EEG spectra aimed at defining the individual biometric pattern. The developed system⁵ was tested on the EEG data recorded from 82 subjects and provided accuracy of 90%. Later studies^{33,34} have confirmed that EEG features contain the information related to an individual trait, which is stable over time.

Classification of EEG recordings from 40 subjects was attempted in the study.⁶ Recordings were obtained with 8 electrodes, and subjects were at resting state with eyes open and eyes closed. The EEG signals were represented by the coefficients of AR models. Linear discriminant analysis was employed for classification. AR models of orders from 3 to 21 were compared and the identification accuracy of 82% was reported with 21 coefficients.

Visually evoked EEG potentials recorded from a multi-electrode system have been investigated for biometric identification of 102 subjects.³⁵ The investigators assumed that the frequency powers in the Gamma band (32-48 Hz) represent the informative features for biometric identification. The spectral powers were estimated for each electrode in order to form a feature vector. Biometric identification was performed with Elman neural networks which allowed an accuracy of 98% to be achieved.

The same EEG benchmark has been studied for identification by using a support vector machine. The biometric features were represented by the spectral powers, and an accuracy of 91% was reported.³⁶ The Physionet Motor Movement/Imagery EEG data set (EEG-MMI),³⁷ including recordings made from 109 subjects via a 64-electrode EEG system, has been used in a number of the recent works in biometric recognition and identification. The recordings were made from the subjects during the rest periods with eyes open or eyes closed. The subjects also performed

motor tasks including real and imagined movements of hands and feet.

The accuracy of biometric authentication has been explored on six publicly available data sets, including the EEG-MMI benchmark. The study⁷ compared the authentication accuracy that was achieved by using the spectral features extracted from the conventional frequency bands, as well as by employing different movement-related tasks. A Bayes classifier was used for the biometric identification. It was reported that the EEG features from a wide range of frequencies were similarly informative for the identification. The analysis of the electrode positions has not revealed their best location in terms of identification accuracy. An accuracy of around 80% was reported for 100 subjects. In a subsequent work by the same authors³⁸ an accuracy of 95% was achieved on the EEG-MMI benchmark by analysing the shape of the power spectrum.

The resting state recordings from the EEG-MMI data set have been used in the experiments.³⁹ The 1-min long recordings were split into 12 5-sec epochs to be represented by the coefficients of AR models of orders given between 5 to 20. The accuracy of biometric identification was reported to be 87.2%.

The EEG-MMI benchmark has been used to research the effects of movement-related tasks on biometric recognition.⁴⁰ The 4-sec long task-related recordings were joined into 30-sec long epochs in order to extract the biometric features by using a wavelet transform. The recognition was performed with linear discriminant analysis, and an accuracy of nearly to 100% was achieved. This approach requires the 30-sec long recordings for the biometric identification, which under certain conditions is not user-friendly and can limit applications.

2.2. Mapping the brain functional connectivity

The brain functional connectivity reflects the interaction between different brain regions, that is observed and explained in one of the existing forms of measuring brain activity.⁴¹ The electrical brain activity can be measured directly and represented by potentials induced on EEG electrodes. The brain activity can be also estimated from the Magnetoencephalogram (MEG)⁴² or functional Magnetic Resonance (fMRI).⁸ The observations are mapped onto the “connectome” that represents her-

itable individual differences in the brain organisation,^{11,12,43} which makes the connectivity-based approach promising for biometric purposes.

In the study⁸ the individual differences in a connectome were defined as “connectotype” analogously to genotype. For estimating the connectome from fMRI data, AR modelling was employed to predict activity of a given brain Region of Interest (ROI) as a weighted sum of the activity of its neighbouring regions.

The neural activity in fMRI data is represented by the Blood-Oxygen-Level Dependent (BOLD) signal which reflects the changes in blood oxygen saturation associated with activation of neurons. The BOLD signal measured in a given ROI is correlated with EEG recorded from an electrode.⁴⁴ As the consequence, the techniques developed for analysis of BOLD signals can be transferable to EEG feature extraction.

The early attempts^{41,45} to map the connectivity used the correlation estimated between the regions of interest, which can be represented by EEG or fMRI signals. The correlation was estimated between pairs of the regions of interests.

Later the connectivity mapping was considered in the frequency domain in terms of coherency which reflects the linear relationship between a pair of EEG channels at a given frequency. However it was found that the interpretation of the coherence is affected by the volume conduction which leads to erroneous estimation of the coherence.⁴⁶

A study was proposed to estimate the connectivity in terms of phase synchrony between EEG electrodes, which is known as Phase Locking Value (PLV). The PLV is based on the observation that during a cognitive process groups of neurons generate the EEG signals with a precise phase synchrony. The observed synchronies between brain regions were capable of reflecting the functional connectivity, whilst the local synchronies observed within a given region were related to the volume conduction. Being influenced by the volume conduction, the neighbouring electrodes receive an EEG component with a similar electrical potential, which affects the mapping of the functional connectivity.⁴⁷

Another approach¹⁷ to mapping the functional connectivity in terms of phase synchronisation was by estimating an imaginary part of the coherence. The imaginary part captures the information about

phase differences of two signals while the real part of coherence (or ordinary coherence) is sensitive to signal magnitude. It was observed that the EEG signals recorded at the scalp electrodes do not have phase differences with the underlying source activity induced by the volume conduction. On the other hand, the presence of phase differences cannot be explained by the influence of the volume conduction. The real and imaginary parts of coherence were explored, and it was found that the imaginary part can reflect the interaction as a biometric feature related to the functional connectivity.

Although the imaginary part of coherence was promising for estimating the true interactions between brain regions,¹⁷ this approach has the following two limitations. First, to achieve reliable results long EEG sessions are required. Second, the coherence is defined under the unrealistic assumption that the phase difference between two signals is constant.¹⁹

An improved mapping of phase interactions was proposed and defined as the Phase Lag Index (PLI).¹⁸ The PLI maps the asymmetry of the distribution of phase differences between two signals. The distribution is expected to be uniform if there is no synchronisation between the two signals, and it becomes asymmetrical if phase synchrony is observed. It was concluded that the PLI is more robust to the volume conduction than the coherence-based mapping.

2.3. *Connectivity-based features*

Connectivity-based features¹⁴ have been proposed for biometric recognition by estimating the coherence between two frontal electrodes. The new features taken along with the estimates of conventional power spectral density have provided an identification accuracy of 98.1% for 51 subjects, using 1-min long EEG recordings.

Another study¹⁵ proposed mapping the coherence along with the standard spectral features for biometric recognition using the resting-state 1-min long EEG recordings from the EEG-MMI benchmark. An identification accuracy of 100% was reported. The accuracy is likely to be overestimated because there was no evidence that the validation subsets, which were used for feature selection, were different from those which were used for estimating the performance. The resting-state data were taken

from a continuous 1-min session, which cannot represent the natural interstate variations. The influence of the volume conduction was not considered within the proposed method.

As described in the previous section, the PLI-based features can be used for extracting features robust to the influence of the volume conduction. Such a method⁹ has been proposed for extracting the connectivity-based features calculated for each pair of EEG electrodes. The generated features were then used for estimating so-called network centrality within which the importance of electrodes is evaluated for identification of the connectome patterns.

Recently,¹⁰ PLV and PLI based features have been compared on the EEG-MMI benchmark. It was shown that the PLV features require EEG signals to be orthogonalised by using linear regression analysis in order to remove the influence of the volume conduction. An accuracy close to 90% was reported for recognising the resting-state EEG of 109 subjects. It was shown that the recognition accuracies of the PLV and PLI features were comparable.

3. Problem Statement

In this section we introduce definitions and then consider the connectotype-based approach to EEG-based identification and state the research problem. As discussed above, the brain functional connectivity, which is individual for a person, can be analysed when the effect of volume conduction is reduced to a minimum. There are three main approaches to achieve this goal: the first and second are based on decorrelation and orthogonalisation of signals respectively, whilst the third estimates phase shifts between multi-electrode signals, which make biometric features robust to the volume conduction. Advantages and limitations of this approach are discussed at the end of the section.

3.1. Volume conduction and brain functional connectivity

The brain activity that is related to a task (e.g. perceptual, motor or mental) performed by a subject, as observed via a multi-electrode EEG system, from the neurophysiologic point of view has two projections or components caused by (i) the volume conduction and (ii) the brain functional connectivity. The first projection is determined by the electrical

conduction of the brain volume, and so causes trivial correlations between EEG electrodes. The signals of this component have similar waveforms and phases over multiple EEG electrodes and so are not related to a task of biometric test, although volume conduction makes a contribution to biometric information on the individual brain morphology.¹⁵ However during biometric test the influence of the brain conduction is trivial and has to be removed from the biometric analysis.

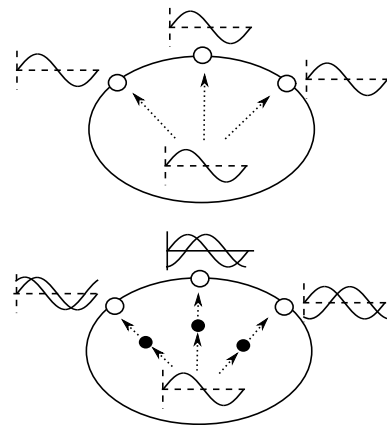


Figure 1. Volume conduction and brain functional connectivity. The upper plot: a source of the neural activity generates a sine wave received by the sensors placed on the scalp. The component caused by the volume conduction is recorded without a phase delay. The lower plot: the sine wave passing different network connections is received with a phase difference on the electrodes.

The second component is caused by the brain functional connectivity that defines the brain “connectotype” of a person. The sources of this EEG component have a unique anatomical localisation and so induce electrode potentials with phase delays.^{48, 49} Fig. 1 illustrates this phenomenon. Being generated from different brain areas, the connectotype components induce the potentials which are correlated with the personal brain functional connectivity and so represent a unique biometric “brain print”.

The EEG component caused by the volume conduction, which is not correlated with the biometric brain print, can be modelled and then removed by orthogonalising the EEG signals received from the electrodes.

3.2. Decorrelation using source derivation

To reduce the effect of brain conduction in a multi-electrode EEG system, Hjorth⁵⁰ has suggested a source derivation method employing the Laplace operator to approximate a distribution of the potential field on the surface of a conductive volume. The method aims to estimate the potential ψ_s of the source component as follows:

$$\psi_s = -(d_x^2 + d_y^2)\psi_{xy},$$

where d_x^2 and d_y^2 are the operators that define the 2nd order differences along x and y coordinates of the surface underlying the scalp electrodes, and ψ_{xy} is the distribution of potentials over the x and y coordinates.

Consider a 5-electrode placement with the central electrode at a position P_0 and the four surrounding electrodes at positions P_1, \dots, P_4 along the x and y coordinates. In particular, the electrode positions P_3, P_0, P_1 are placed along the x coordinate, whilst the others P_4, P_0, P_2 are placed along the y coordinate. In this framework, the operators d_x^2 and d_y^2 are defined as follows: $d_x^2 = (\psi_1 - \psi_0) - (\psi_0 - \psi_3)$ and $d_y^2 = (\psi_2 - \psi_0) - (\psi_0 - \psi_4)$.

Finally the potential of interest ψ_s is written as a superposition of four radial bipolar derivations around the central electrode:

$$\psi_s = (\psi_0 - \psi_1) + (\psi_0 - \psi_2) + (\psi_0 - \psi_3) + (\psi_0 - \psi_4). \quad (1)$$

Fig. 2 shows the block diagram of the source derivation method implementing the above Eq. 1.

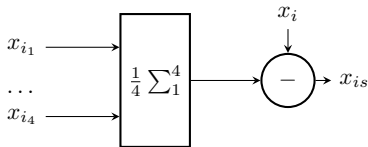


Figure 2. Block diagram of the source derivation method implementing Eq. 1. Here x_{i1}, \dots, x_{i4} are the signals from four electrodes neighbouring to the central electrode, x_i is the signal from the central electrode, and x_{is} the derived component of the signal.

The common reference derivation for M electrodes has a similar form that can be written as

$$\psi_s = \psi_0 - \frac{1}{M} \sum_{i=1}^M \psi_i. \quad (2)$$

However the source derivation method, implementing Eq. 1, excludes superimposed contributions from sources located outside of neighbouring areas in a multi-electrode system. As a result, the source derivation has improved the separation by a factor of 2 compared with the common reference implementing Eq. 2 in terms of correlation caused by the brain conduction. Correspondingly, the estimated source component will have a greater signal-to-noise ratio.

An attempt⁵¹ to improve the source derivation resolution within the 10-20 system of electrode placement has been made by applying weights $1/(1 + R^2)$ to contributions of the electrodes, where R stands for the distances from the electrode position to the surface element. However this technique has produced only a minor improvement of the resolution.

Analysing the source derivation method, Hjorth⁵⁰ has found that correlations between neighbouring and central electrodes have different signs: nearest electrodes often have a positive correlation, whilst more remote electrodes often have a negative correlation. He has concluded that this can explain why the common reference technique (Eq. 2) applied within the standard 10-20 electrode system cannot improve the source separation resolution.

Further attempt⁵² to obtain the optimal resolution has been made by adaptation of the derivation technique using covariances in the potential field y observed on the scalp electrodes. According to the method, the brain activity is represented by the signals $s = (s_1, \dots, s_M)$ referenced to the common average $x = ([\mathbf{I}] - 1/M[\mathbf{1}])s$, where $[\mathbf{I}]$ is the identity matrix and $[\mathbf{1}]$ the unit matrix.

Resultant superposition in the potentials y depends on the transfer function matrix \mathbf{A} : $y = \mathbf{A}x$, where x are the intrinsic data subject to convolution by the matrix \mathbf{A} .

Independence and equal variances of x are the necessary conditions to relate the covariance matrix \mathbf{C} to \mathbf{A} : $\mathbf{C} = yy^T = \mathbf{A}x(\mathbf{A}x)^T = \sigma_x^2 \mathbf{A}\mathbf{A}^T$, that make it possible to identify the matrix \mathbf{A} : $\mathbf{A} = \sqrt{\mathbf{C}}$.

The correlations which are caused by off-diagonal elements of \mathbf{A} cannot be directly minimised by the inverse of \mathbf{A} without a term $1/M[\mathbf{1}]$, addition of which finally brings the solution to the problem:

$$\mathbf{B} = (\mathbf{A} + 1/M[\mathbf{1}])^{-1}. \quad (3)$$

This solution nevertheless can be numerically unstable, thus affecting the desired decorrelation of

the source $z : z = \mathbf{B}\mathbf{y}$. In practice, the adaptive source derivation method has produced a minor improvement.⁵²

3.3. Orthogonalisation using regression modelling

In EEG, MEG, and fMRI modalities of brain activity imaging, the connectomes can be represented by a residual component of the signal predicted with an AR model that has been obtained for each person during the enrolment (training) sessions.⁸ The residuals r are mathematically defined as components which are orthogonal to the predicted signal \hat{y} . Being orthogonal, the residuals are affected by the trivial correlation to a lesser degree.^{19,20} The common idea of these methods can be outlined as follows.

Signals $\{x_i(t)\}_{i=1}^M, t = 1, \dots, T$, are obtained from an M -channel system, as part of a biometric test that is performed by a person during the enrolment and identification sessions in the form of motor or mental task. Here T is the number of samples in recording of a task which is used for the enrolment and identification.

The signal $x(t)$ is represented by an AR model:

$$x(t) = \sum_{j=1}^P \alpha_j x(t-j) + r(t), \quad (4)$$

where α_j are the coefficients of the AR model, P is the order of the model, and r is the residual component which is associated with the connectome.

Given the EEG samples for each enrolment and identification session, the desired coefficient vectors $\hat{\boldsymbol{\alpha}}$ are defined as follows

$$\hat{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{\alpha}} S(\boldsymbol{\alpha}), \quad (5)$$

where $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_P)^\top$.

Here $S(\boldsymbol{\alpha})$ is the error function calculated for the residuals r :

$$S(\boldsymbol{\alpha}) = \sum_{t=P+1}^T r_t^2 = \sum_{t=P+1}^T \left[x(t) - \sum_{j=1}^P \alpha_j x(t-j) \right]^2.$$

In the matrix form the above problem can be rewritten as follows

$$S(\boldsymbol{\alpha}) = \|\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}\|. \quad (6)$$

Here $\mathbf{y} = (x_T, \dots, x_{P+1})^\top$, and \mathbf{X} is the data matrix:

$$\mathbf{X} = \begin{bmatrix} x_{T-1} & x_{T-2} & \dots & x_{T-P} \\ x_{T-2} & x_{T-3} & \dots & x_{T-P-1} \\ \dots & \dots & \dots & \dots \\ x_P & x_{P-1} & \dots & x_1 \end{bmatrix}.$$

The above minimisation problem (Eq. 5) has a unique solution $\hat{\boldsymbol{\alpha}}$ when all P columns of \mathbf{X} are linearly independent. In this case the solution can be found by solving the normal equations:

$$(\mathbf{X}^\top \mathbf{X})\hat{\boldsymbol{\alpha}} = \mathbf{X}^\top \mathbf{y},$$

that finally allows us to calculate the coefficient vector $\hat{\boldsymbol{\alpha}}$ using the inverse operator $(\cdot)^{-1}$:

$$\hat{\boldsymbol{\alpha}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}. \quad (7)$$

The residual vector \mathbf{r} is written as follows:

$$\mathbf{r} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\alpha}}. \quad (8)$$

According to the method,⁸ the AR models of order $P = 5$, described by Eq. 4, are built in order to remove the trivial correlation caused by the dispersion of signals in each channel $m = 1, \dots, M$ during the enrolment sessions. The correlations are removed from the data of each person.

The residual vectors $\mathbf{r}_m, m = 1, \dots, M$, are calculated using Eq. 8 are then used for analysis of the brain functional connectivity. Then the residuals are arranged for the i th lag (frame) $t = P + 1, \dots, T$ as follows $\mathbf{r}_t = (r_{1,t}, \dots, r_{M,t})^\top$ in order to model the neural interaction between the electrodes (ROIs), where P is the number of lags.

The interaction is linearly modelled as follows:

$$\hat{\mathbf{r}}_t = \sum_{j=0}^P \mathbf{B}_{i-j} \mathbf{r}_{i-j}, \quad (9)$$

where \mathbf{B}_j is a $M \times M$ matrix:

$$\mathbf{B}_j = \begin{bmatrix} 0 & b_{1,2,j} & \dots & b_{1,M,j} \\ b_{2,1,j} & 0 & \dots & b_{2,M,j} \\ \dots & \dots & \dots & \dots \\ b_{M,1,j} & b_{M,2,j} & \dots & 0 \end{bmatrix}. \quad (10)$$

The coefficients of \mathbf{B}_j are determined by using a Pseudo-Inverse (PINV) technique, which similarly to the technique based on normal equations (Eq. 7) provides a unique solution. The PINV-based solution is numerically stable when the numbers of rows and columns are equal, and M is large ($M = 184$), whilst

a technique based on the normal equations requires a larger number of rows than columns.

Within the above method, the coefficient vectors $\boldsymbol{\alpha}^{(l)}$ and coefficients matrices $\mathbf{B}^{(l)}$ are calculated for each of L enrolled persons $l = 1, \dots, L$ and then saved for an identification session.

During identification the test data received from a person via the M channels are used to calculate the residual vectors \mathbf{r} given the vectors $\boldsymbol{\alpha}$ and \mathbf{B} obtained for each enrolled persons.

Then it is assumed that the residual vectors \mathbf{r} (Eq. 8) and $\hat{\mathbf{r}}$ (Eq. 9), calculated for the test of the l th person with the parameters $\boldsymbol{\alpha}^{(i)}$ and $\mathbf{B}^{(i)}$, are statistically similar if $i = l$ and significantly different if $i \neq l$, where $i = 1, \dots, L$.

The similarity between the residuals \mathbf{r} and $\hat{\mathbf{r}}$ is estimated over the test by using a correlation coefficient r_{xy} . In Pearson's definition, the correlation coefficient r_{xy} is estimated between the samples x and y .

The correlation coefficients r_{xy} are calculated for each channel and then the mean correlation over all M channels is estimated for each candidate l .

The person l^* , whose mean correlation coefficient \hat{r}_{xy} is maximal, is identified to be with the best matching brain print:

$$l^* = \arg \max_{1 \leq l \leq L} (\hat{r}_{xy}^{(l)}),$$

The MRI-based identification method⁸ has been tested on a benchmark set including 27 persons each of whom is represented by a trial including a few samples (frames). The training and test frames have been randomly selected from the same trial. The identification accuracy has been evaluated on the test frames which can be correlated with the training frames, and so the reported performance 100% can be overestimated. The second trials have been taken from only 5 persons, which cannot be sufficient.

3.4. Phase-based approach

According to^{17, 18, 47} the brain connectivity influences the phase delays φ in the EEG signals received from different brain regions via a multi-electrode system. The neural activity in a local area passing different network connections induces a distinct electrical potential on EEG electrodes. Because the sources are located in different areas, the induced EEG signals have different phases ϕ_i so that the phase delays φ_{ij}

between electrodes i and j are

$$\varphi_{ij} = \phi_i - \phi_j, i \neq j = 1, \dots, M. \quad (11)$$

The phases ϕ_i are estimated instantaneously.¹⁸ Their estimates contain noise components assumed to have a Gaussian distribution with variance v_i^2 . The resultant phase delays φ_{ij} will be therefore affected by noise with a larger variance $v_{ij}^2 = v_i^2 + v_j^2$.

The delays φ_{ij} are estimated for each pair of electrodes, and their number is $M(M-1)/2$. This can be large for multi-electrode systems, for example when $M = 64$, the number becomes 2016. The delays are estimated for each frequency that can be present in the EEG signal.

3.5. Limitations of the approaches

The analysis of the above EEG-based biometric methods reveals the following limitations.

(1) Some studies have reported identification performances which are statistically overestimated because of the following reasons: (i) the test and training samples (frames) are taken from the same trial and (ii) the test data include a small number of participants. However reliable estimates can only be obtained on test data having a sufficient number of samples not included in the training data set.

(2) The EEG-based trials have been reported to be 30-sec, and MRI-based methods have 2-min trials. For some biometric applications such long trials cannot be considered as user-friendly.

(3) Both EEG-based and MRI-based methods deal with multidimensional data. The analysis of such high-dimensional data is connected with general problems of calculating a pseudo-inverse of the matrix \mathbf{B} (Eq. 10), which has to be capable of providing a numerically stable solution.

(4) The interaction with the historical lags in the residual vector $\hat{\mathbf{r}}$ (Eq. 9) is not modelled. The inclusion of new lags, $j = 1, \dots, P$, will therefore cause a problem of computing a pseudo-inverse of matrices \mathbf{B}_j .

(5) Being affected by the noise and artefacts, some of the delays φ_{ij} in Eq. 11 as biometric features can be irrelevant to the personal brain print and so have to be removed in order to achieve reliable identification. These above problems limit the reliability of biometric identification.

4. Proposed Method

The above findings motivate us to consider ways to overcome the existing limitations and develop a new approach capable of improving the accuracy and reliability of EEG-based person identification. The new method should also be user-friendly and provide the identification result using biometric tests of a reasonably short duration.

4.1. Statement 1

We found that the limitations discussed in the above section can be overcome with a new technique which can (i) use a smaller set of features than that provided by the pairwise interactions and (ii) be capable of reducing the intensity of noise components present in φ .

Both these requirements are satisfied by using a “virtual” (in contrast to a physical electrode that is used as the reference) baseline instead of the above pairwise-electrode interaction. The idea behind the proposed approach is to use the neighbouring electrodes for modelling the baseline, which is required for estimating the brain connectivity represented by the phase delays φ .

Statement 1: Residuals of Autoregressive modelling of EEG signals referenced to the proposed “virtual” baseline are correlated with brain functional connectivity.

To support Statement 1, let us assume a model that can predict the EEG signal observed at the i th electrode using the signals obtained from the other electrodes so that the predicted signals form the “virtual” baseline observed with a phase $\hat{\phi}_i$. Then the delay φ_i for the i th electrode can be estimated as follows:

$$\varphi_i = \phi_i - \hat{\phi}_i, i = 1, \dots, M.$$

Here in contrast to the delays φ_{ij} calculated for the pairwise interactions described by Eq. 11, the above features are estimated only for M electrodes.

The use of an AR model of an order P for predicting the EEG signals reduces the noise component that contaminates the phases ϕ_j in Eq. 11. The total sum of the absolute values of φ_{ij} over all T predicted samples is proportional to the total sum of the resid-

ual $|r_i|$ at the i th electrode:

$$\sum_{t=P+1}^T |\varphi_{it}| \rightarrow \sum_{t=P+1}^T |r_{it}|. \quad (12)$$

The phases mapped in this way onto the residuals are used in our experiments as biometric features capable of making important contributions to the identification.

The distribution of the phases φ over the multiple electrodes is the unique biometric pattern of interest. In this connection, the research¹⁸ has demonstrated that the brain connectivity can be estimated by evaluating the shape of a distribution of phase differences between two signals. This distribution is uniform when there are no interactions between the signals, and the interactions exist if the distribution is asymmetrical.

4.2. Statement 2

It is important to note that improvement of the accuracy of modelling the volume conduction by adding new inputs (channels) to the AR, described by Eq. 4, explicitly increases the dimensionality of the coefficient vector α . An increase in the dimensionality of an AR model will cause problems of finding a numerically stable solution capable of minimising the error function $S(\alpha)$ in Eq. 5 and providing the vector $\hat{\alpha}$ by using the inverse operator used in Eq. 7.

Given a limited set of the enrolment data, reliable estimation of AR model parameters can be achieved with the regularisation criterion,^{22,26} Δ , that is evaluated on the disjoint subsets \mathcal{A} and \mathcal{B} of the enrolment data set \mathcal{W} :

$$\mathcal{W} = \mathcal{A} \cup \mathcal{B}, \mathcal{B} \not\subset \mathcal{A}.$$

The subsets \mathcal{A} and \mathcal{B} are assigned for fitting and validation of AR models, respectively. The subset \mathcal{A} is used for fitting the coefficient vector $\alpha_{\mathcal{A}}$ of an AR model $f(\mathbf{X}_{\mathcal{A}}; \alpha_{\mathcal{A}}, \mathcal{X})$ given with a structure vector of the input variables $\mathcal{X} = (x_{i_1}, \dots, x_{i_m})$, here $1 \leq m \leq M$.

The desired regularisation is achieved by evaluating the residual Δ of the model $f(\cdot; \alpha_{\mathcal{A}}, \mathcal{X})$ on the entire enrolment data including the validation subset \mathcal{B} :

$$\Delta = \|f(\mathbf{X}; \alpha_{\mathcal{A}}, \mathcal{X}) - \hat{\mathbf{y}}\|. \quad (13)$$

In the absence of knowledge of the structure vector \mathcal{X} , a reasonable set of K vectors ($\mathcal{X}_1, \dots, \mathcal{X}_K$) is

generated so that the residuals of AR models are arranged as follows:

$$\Delta_{i_1} \leq \Delta_{i_2} \leq \dots \leq \Delta_{i_K}.$$

The model $f^*(\cdot; \boldsymbol{\alpha}, \mathcal{X}_{i_1})$ with the structure vector \mathcal{X}_{i_1} , which provides the lowest residual Δ_{i_1} , is selected and assigned to be “optimal”.

A family of GMDH algorithms^{22, 26, 27} can learn the model of interest under the above conditions, namely (i) when the enrolment data \mathbf{X} are represented with a small number of instances or under-determined ($N < M$), and (ii) when the structural vector \mathcal{X} is unknown.

Statement 2: GMDH can learn the “optimal” structure of Autoregressive model from given enrolment data.

To support Statement 2, let the signals $x^{(i)}$, $i = 1, \dots, M$, be the result of interaction between M electrodes, which can be modelled by AR models:

$$x^{(i)}(k) = \sum_{m=1}^M \beta_m^{(i)} \sum_{j=1}^{P_i} \alpha_j^{(i)} x^{(i)}(k-j) + r^{(i)}(k), \quad (14)$$

where $\beta^{(i)} = (\beta_1, \dots, \beta_M)$ are the vectors of coefficients of an electrode i with $\beta_m^{(i)} = 0$ when $m = i$, $P_i \in \{1, P\}$ are the orders of AR models at the electrode i , which are unknown but limited to the maximum P , and $r^{(i)}$ are the residual of the electrode i , which is associated with the brain connectivity that is the subject of biometric estimation.

The delayed components of the signal, $x^{(i)}(k-j)$, are included in the above Eq. 14 to model a contribution of the brain activity which is caused by the functional connectivity.

The AR models will therefore include the electrode coefficient vector $\boldsymbol{\beta} = (\beta_1, \dots, \beta_M)$ so that the residual $\Delta^{(i)}$ at each channel is defined:

$$\Delta^{(i)} = \|f_i(\mathbf{X}; \boldsymbol{\beta}, \boldsymbol{\alpha}, \mathcal{X}) - \hat{\mathbf{y}}\|, \quad (15)$$

where the underscripts \mathcal{A} for the above $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ are omitted for simplicity.

Therefore, in the multichannel framework the solution to the minimisation problem needs to be found for an extended vector of the parameters $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$ and \mathcal{X} :

$$(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \hat{\mathcal{X}})^{(i)} = \arg \min_{\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathcal{X}} \Delta^{(i)}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathcal{X}). \quad (16)$$

The block diagram of the proposed method implementing Eq. 14 is shown in Fig. 3.

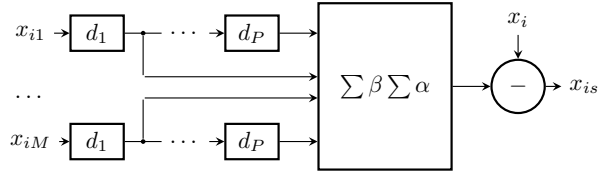


Figure 3. Block diagram of the proposed method implementing Eq. 14. Here d_1, \dots, d_P are the delay units, x_i is the central electrode, and x_{is} the orthogonalised component of the signal.

Because of the limitations discussed in the subsection 3.5, finding the reliable solution, based on an extended parameter vector $(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \hat{\mathcal{X}})$, will be undertaken within the GMDH framework introduced in section 4.2. The proposed GMDH algorithm is discussed in the next section.

4.3. GMDH polynomial algorithms

Having outlined the main idea of GMDH algorithms in the section 4.2, here we consider a family of GMDH-type polynomial networks and then outline the limitations of the existing GMDH approaches.

4.3.1. Description

Given a polynomial reference function with arguments z_1, \dots, z_μ , GMDH algorithms can approximate an arbitrary polynomial of interest with M arguments. Within GMDH, the approximation accuracy is estimated by using the regularisation criterion Δ , which is described by Eq. 13 for given data partitions \mathcal{A} and \mathcal{B} .

For example, given an order of 3, a polynomial of interest can be described as

$$y = w_0 + \sum_{i=1}^M w_i z_i + \sum_{i=1}^M \sum_{j=1}^M w_{ij} z_i z_j + \sum_{i=1}^M \sum_{j=1}^M \sum_{n=1}^M w_{ijn} z_i z_j z_n,$$

where w_i , w_{ij} and w_{ijn} are the coefficients for the terms of order 1, 2 and 3, respectively.

A reference function g described by a short-term polynomial of 2 arguments z_1, z_2 can be considered as a neuron activation function. In this notation, GMDH algorithms build ANNs which can be learnt from data without prior information about ANN structure.^{22, 28}

In the case of 2 arguments, an activation function g with the coefficient vector \mathbf{w} is written as follows:

$$y = g(z_1, z_2; \mathbf{w}).$$

We can therefore define the polynomial activation functions of the 1st and 2nd orders, the most common of which are:

$$y_1 = w_0 + w_1 z_1 + w_2 z_2,$$

$$y_2 = w_0 + w_1 z_1 + w_2 z_2 + w_{12} z_1 z_2,$$

$$y_3 = w_0 + w_1 z_1 + w_2 z_2 + w_{12} z_1 z_2 + w_{11} z_1^2 + w_{22} z_2^2.$$

The 1st order polynomial y_i is used for modelling of linear systems. The 2nd order polynomials y_2 and y_3 are used for modelling of nonlinear systems.

Having the training data $[\mathbf{X}; \hat{\mathbf{y}}]$, including a $N \times M$ -matrix of the enrolment data \mathbf{X} and the target vector $\hat{\mathbf{y}}$, we could set a polynomial order and then define a corresponding activation function g .

Below we provide details of our algorithm modified for the experiments with the EEG biometric data. Using a given activation function, GMDH generates new neurons which are fitted to the training data $[\mathbf{X}; \hat{\mathbf{y}}]$. The generated candidate-neurons are then evaluated in terms of the regularity error Δ that is estimated on the validation data. Such a construction enables GMDH to select neurons with the best generalisation ability.

The neurons, which have been selected at the layer r , generate features for neurons at a new layer $r + 1$. The number of layers r is consequently increased, and the network is growing, whilst its generalisation ability is improved. During this phase the values of regularity error Δ calculated for layers $1, 2, \dots, r$ tend to decrease. Once at the layer $r^* = r + 1$, the network becomes overcomplicated, its error $\Delta^{(r^*)}$ will increase according to the regularisation criterion. Having reached this phase, the GMDH algorithm is terminated as illustrated on Fig. 4.

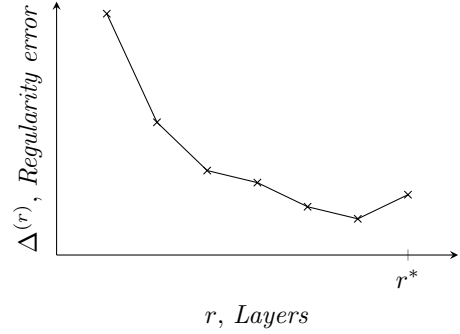


Figure 4. Regularity errors $\Delta^{(r)}$ over layers $r = 1, 2, \dots, r^*$. The algorithm is terminated at the layer $r^* = r + 1$ when $\Delta^{(r+1)} > \Delta^{(r)}$.

4.3.2. Problems of learning GMDH-type ANN

The candidate-neurons, which have been generated at a layer r , are arranged by their regularity errors Δ in ascending order to be selected for the next layer. The first F neurons with the lowest errors are selected as follows:

$$\Delta_{i_1} \leq \dots \leq \Delta_{i_F} \leq \dots \leq \Delta_{i_K}. \quad (17)$$

where i_1, \dots, i_K are the indexes and K is the number of the candidate-neurons in a layer r .

In conventional GMDH algorithms, the “optimal” number F is typically set to be around $0.4K$ and refined in experiments. GMDH selects the same number of neurons in each layer $r = 1, 2, \dots, r^*$, which under certain conditions provides the reliable results.²² However such a selection is not efficient when a network with a large number of layers tends to “degrade” in terms of the diversity of neuron outputs. The selected neurons generate correlated outputs which are not capable of making a distinctive contribution.

4.4. Proposed GMDH algorithm

The analysis of the above limitations gives new insights into the problem of GMDH and we describe an algorithm and its implementation.

4.4.1. Description

An efficient solution to the above problem is based on removal of the correlated outputs. The modified algorithm is terminated at a layer r^* when the number of independent outputs is below a given threshold.

Using a given activation function $g(z; \mathbf{w})$, the algorithm generates a feature vector $z = (x_{i_1}, x_{i_2})$ for neurons at the 1st layer ($r = 1$), where $i_1 \neq i_2, i_1 = 1, \dots, M$. The number of generated neurons is $K_r = M(M-1)/2$ that is defined by the number of pairwise combinations for M input variables at the 1st layer.

The outputs \mathbf{y}_i of neurons at layer $r = 1$ are written as:

$$\mathbf{y}_i^{(r)} = g(z; \hat{\mathbf{w}}_i^{(r)}), i = 1, \dots, K_r. \quad (18)$$

Given a set \mathcal{A} of the training instances, the coefficient vectors $\hat{\mathbf{w}}^{(r)}$ are fitted to the data $[\mathbf{X}(\mathcal{A}); \hat{\mathbf{y}}(\mathcal{A})]$ by solving the normal equations. Their solutions are numerically stable as the number of instances included in the set \mathcal{A} is larger than the number of the inputs in the activation function, and the variables in columns i_1, i_2 of the matrix $\mathbf{X}(\mathcal{A}, [i_1, i_2])$ are not correlated. The estimates of the weight coefficients $\hat{\mathbf{w}}$ therefore are:

$$\hat{\mathbf{w}}^{(r)} = [\mathbf{1} \ \mathbf{X}(\mathcal{A}, [i_1, i_2])]^{-1} \hat{\mathbf{y}}(\mathcal{A}), \quad (19)$$

where $\mathbf{1}$ denotes the unit vector.

Having the outputs \mathbf{y}_i on the entire data \mathbf{X} , the regularity errors Δ is calculated as follows:

$$\Delta_i = \|\mathbf{y}_i - \hat{\mathbf{y}}\|. \quad (20)$$

The errors are sorted into ascending order (Eq. 17), and then the first F neurons with lowest errors are selected for the next layer.

As discussed in section 4.3.2, the outputs \mathbf{y}_i can be correlated with outputs of the other neurons generated at a layer r . To avoid this problem, the proposed algorithm excludes the neurons with correlated outputs from the $\mathbf{y}_{i_1}, \dots, \mathbf{y}_{i_F}$ to be selected for the next layer. So the number of selected neurons after decorrelation can vary K_r : $1 \leq K_r \leq F$.

Neurons at the next layers $r + 1$ are generated by applying the function g to an extended data matrix $\mathbf{Z}^{(r)} = [\mathbf{Y}^{(r)}; \mathbf{X}]$, which includes the outputs of the selected neurons $\mathbf{Y}^{(r)} = [\mathbf{y}_{i_1}, \dots, \mathbf{y}_{i_{K_r}}]$, and the input data \mathbf{X} , so that the matrix $\mathbf{Z}^{(r)}$ contains $M_r = M + K_r$ columns, where $M_{r=1} = M$.

Similarly, the coefficient vectors $\hat{\mathbf{w}}^{(r+1)}$ are estimated for the (i_1, i_2) columns of the matrix $\mathbf{Z}^{(r)}$. The outputs of neurons $\mathbf{y}^{(r+1)}$ are then calculated as follows:

$$\mathbf{y}_i^{(r+1)} = g(z_{i_1}^{(r)}, z_{i_2}^{(r)}; \hat{\mathbf{w}}_i^{(r+1)}), i = 1, \dots, K_r, \quad (21)$$

where $i_1 \neq i_2; i_1 = 1, \dots, m_r, i_2 = 1, \dots, M_r$, and $K_r = M_r(M_r - 1)/2$ are the number of pairwise combinations for the M_r columns. Fig. 5 illustrates fitting and selection of neurons $\mathbf{y}_i^{(r)}$ at a layer r .

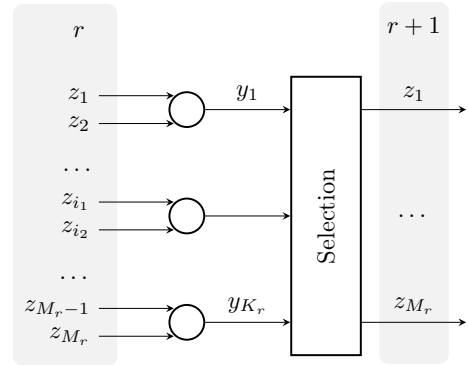


Figure 5. Fitting and selection of neurons $\mathbf{y}_1, \dots, \mathbf{y}_{K_r}$ at a layer $r + 1$ are described by Eq. 21 and Eq. 17 respectively.

The algorithm generates new layers and the network grows while the number of neurons K_r selected at a layer r is larger than a given threshold F_0 . The block diagram of the proposed GMDH algorithm is shown on Fig. 6.

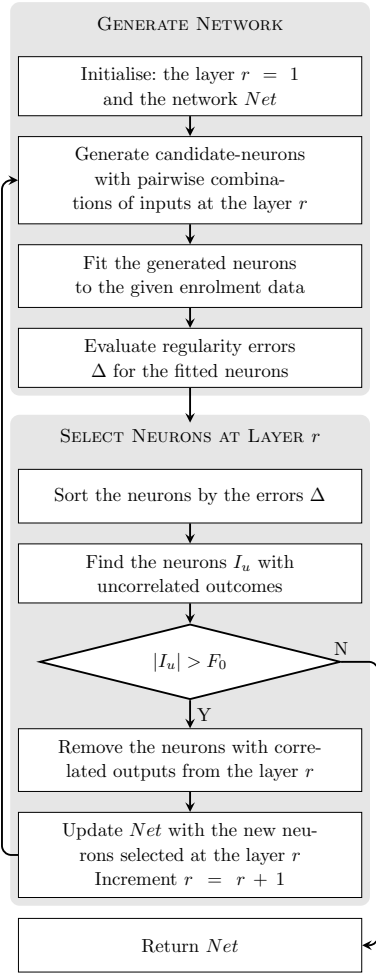


Figure 6. The block diagram of the proposed GMDH algorithm. The candidate-neurons are generated using Eq. 21. Then the neurons are fitted to the given enrolment data according to Eq. 19. The fitted neurons are evaluated in terms of the regularity error Eq. 20. The candidate-neurons are then sorted by using Eq. 17. The neurons, outcomes of which calculated by Eq. 21 are uncorrelated, form a set I_u . The network Net is updated with the new neurons I_u while their number $|I_u|$ exceeds the given minimum F_0 : $|I_u| > F_0$. Finally the Net includes a GMDH-type multilayered network learnt from the given data.

5. EEG Data

The EEG Motor Movement/Imagery Data set (EEG-MMI),³⁷ available from Physionet,⁵³ includes over 1,500 one- and two-min EEG recordings. The recordings were obtained from 109 subjects, and each subject performed the same sets of tasks, which include real and imagined motor movements. The EEG were recorded at 160 Hz sampling rate using 64 electrodes

positioned according to the standard international 10-10 system. The EEG signals are received by the electrodes with numbers 1 to 64 as shown in Fig. 7

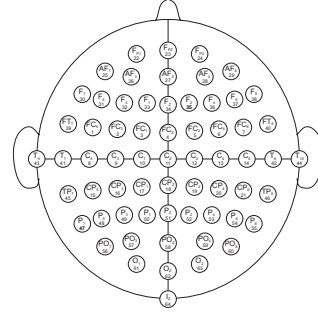


Figure 7. Positions of the 64 electrodes on the scalp.³⁷

Each subject was instructed to relax, and 2-min long resting state EEG recordings were made. After that each subject performed movement and imagery tasks in the response to cues shown on the screen. During the movement tasks when a cue appears on the screen, the subject opens and closes the left or right fist, or both fists or feet. During the imagery tasks the subject is instructed to imagine the movements.

Each task was performed in 2-min long experimental runs, which include on average 15 trials corresponding to the real/imagined movements. Each trial was 4-sec long following after a short resting period. The movements of the left or right fist (or both fists or feet) were assigned randomly with equal probabilities.

It has been shown,¹⁷ that a task determines a specific connectome. However in our research we will study a connectome which is evoked by one type of task, which is the left fist movement, including in total 2,471 trials, on average, 23 trials per subject.

The subjects were instructed to remain still during the recording in order to minimise the occurrence of muscle or electrode movement artefacts. However, the EEG recordings contain some artefacts related to eye movements. We assume that these artefacts are natural and so did not removed them. The raw EEG data were however preprocessed by filtering

slow drifts with frequencies below 0.5 Hz, as well as filtering noise and high-frequency interference above 50 Hz.

6. Experiments

In this section we discuss the experiments which have been run on the EEG-MMI data described in section 5. The aim of these experiments is to compare the proposed and existing methods in terms of identification accuracy and to provide evidences for the statements outlined in section 4.2.

The data used in our experiments include the EEG recordings of the Task 1 type, which have been made from 109 subjects. The average number of trials was 23, and the performances (in terms of identification accuracy on the test trials) were compared within 3-fold cross-validation.

6.1. Comparison of feature extraction methods

The first experiments were run to compare the performances of the proposed and existing feature extraction methods. The proposed method described in the above section 4.2 represents the residuals Δ_i (Eq. 15) as the feature vector $\Delta = (\Delta_1, \dots, \Delta_M)$. According to the statements, the residuals are assumed to be correlated with the functional connectivity of a subject.

The GMDH-type network is trained on the subset \mathcal{A} , whilst the subset \mathcal{B} is used for validation as described in section 4.2. The remaining instances which have not been included in the subsets \mathcal{A} and \mathcal{B} are used for testing the performance of the trained network.

To identify the given L subjects (classes), the GMDH algorithm has been used within the pairwise (round-robin) scheme^{54,55} that requires $L(L-1)/2$ pairs of binary classifiers.

The proposed feature extraction method has been compared with the two conventional methods. The first method⁷ employs the Spectral Power Densities (PSD) which are calculated within 0.5 to 40 Hz, and the second⁸ employs the AR models in order to calculate the residuals and then find the best match using the correlation coefficients.

McNemar’s test, which is used for comparison of classification methods, has shown a statistically significant improvement of the identification accuracy

for the proposed GMDH method. Table 1 shows the performances of the PSD, AR, and GMDH methods in terms of identification accuracy, shown as the mean and the standard deviation over the folds along with the McNemar’s test p -values.

Table 1. Performances of the conventional PSD and AR methods *versus* the proposed GMDH method.

#	Method	Performance, %	p -value
1	PSD	84.5 ± 1.8	< 0.001
2	AR	93.1 ± 1.2	< 0.01
3	GMDH	100.0 ± 0.0	

We can see that the proposed GMDH method provides the most accurate (100%) identification and significantly outperforms the existing PSD and AR methods on this benchmark data using 4-sec EEG segments for the biometric test. In this regard, it is important to note that the study⁴⁰ has reported a 99% accuracy on the EEG-MMI Task 1 data, which include 108 subjects, using a much longer 30-sec duration of EEG segments, which can limit cases of user-friendly applications.

6.2. Comparison of classification schemes

In the second experiments we compared the performances of the classification schemes using the residual-based feature vector Δ . Table 2 shows the experimental results.

The conventional ANNs trained within the multiclass scheme (MANN) have provided the best performance (83.4%) with around 90 neurons in the hidden layer. Within the pairwise scheme (PANN), the ANNs with around 8 neurons have provided a better performance, on average 95.9%. The Support Vector Machine (SVM), which has been also trained in the pairwise scheme, provided the best performance at 99%.

Table 2. Performances of MANN, PANN, and SVM using the features extracted by the proposed GMDH method.

#	Algorithm	Performance, %
1	MANN	83.4 ± 2.8
2	PANN	95.9 ± 1.3
3	SVM	99.1 ± 0.6

The above results show that the new features Δ , extracted by the proposed method, carry the information about connectotypes which can be identified within the main classification schemes.

6.3. Influence of muscle artefacts

Fig. 8 shows the EEG signals of a 4-sec trial along with the predicted and residual signals. The upper plot shows the EEG signal (in Red) recorded from electrode 2 (FC3), and the signals from its 4 neighbouring electrodes, which have been selected by the proposed method. The signals from the neighbouring electrodes were used to predict the signals from electrode 2 in order to generate the “virtual” baseline within the proposed method. The predicted signals, which form the baseline, are shown on the middle plot (in Blue) along with the actual signal (in Red) from electrode 2. The residual signal is shown on the lower plot.

We can observe that the residual signal is increased, for example, at around 600, 3000 and 3400 ms, when a phase difference appears between the actual and predicted signals. Eq. 12 explains this observation for the proposed method. These peaks are not correlated with the ocular artefacts occurring at around 900 and 2300 ms.

6.4. Examples of brain prints

Besides the identification accuracy achieved in our experiments on the EEG-MMI data including 109 participants and shown in Table 1, Fig. 9 illustrates examples of brain prints of 5 persons taken from this benchmark. Each person here is represented by a sequence of 5 brain prints which have been learned from a 4-sec EEG trial.

The brain prints are represented by the residuals which were calculated for all 64 electrodes placed on the scalp as described in section 5. The residual values range from low (in Blue) to high (in Yellow).

Analysing the colour distributions over all the electrodes, we find that the patterns for each person are visually different. We can conclude that the residuals, visualised in colour, carry information about brain prints or connectotypes of persons. In contrast, the patterns are similar within a sequence of 5 trials taken from a person.

Fig. 10 shows the brain prints averaged over all the trials for each subject.

The large residual value at channel 25 can be caused by the activity of the forehead muscles. The large values are observed for persons 1, 2 and 4 across the trials. However this activity is not observed for persons 3 and 5. We therefore can accept that such a muscle activity is individual⁵⁶ and so associated with a biometric feature. This allows us to conclude that muscle artefacts influencing residuals are part of the biometric trait of a person.

6.5. Examples of Power Spectra Distributions of residuals

The residuals can be represented by the power spectral distribution (PSD) calculated within the standard frequency bands: Delta 0-4 Hz, Theta 4-8 Hz; Alpha 8-12 Hz, Beta 12-30 Hz and Gamma 30-50 Hz. The variations in PSD over the scalp electrodes are individual and can be visualised as an EEG-based trait.

Fig. 11 shows the two traits extracted from EEG of two persons recorded in two trials. Although PSD vary over the scalp electrodes, some regularity is visually recognised in the traits extracted for a person. For example, observing the trait of person 1, we see that the high powers are observed in the Beta and Gamma bands (Green and Yellow bars) at the frontal and temporal electrodes: 22-26, 28-31, 38, 40, and 42. The trait of person 2 shows PSD with the low-frequency components in the Delta and Theta bands (Dark and Light Blue bars) at the frontal, temporal, and central electrodes: 1-5, and 15-21. The two traits of a person are similar in both trials. For example, the similarity is observed between electrodes 22-38 in the two trials undertaken by person 1.

7. Discussion

In this section we discuss the main results in terms of the accuracy of the proposed method for biometric

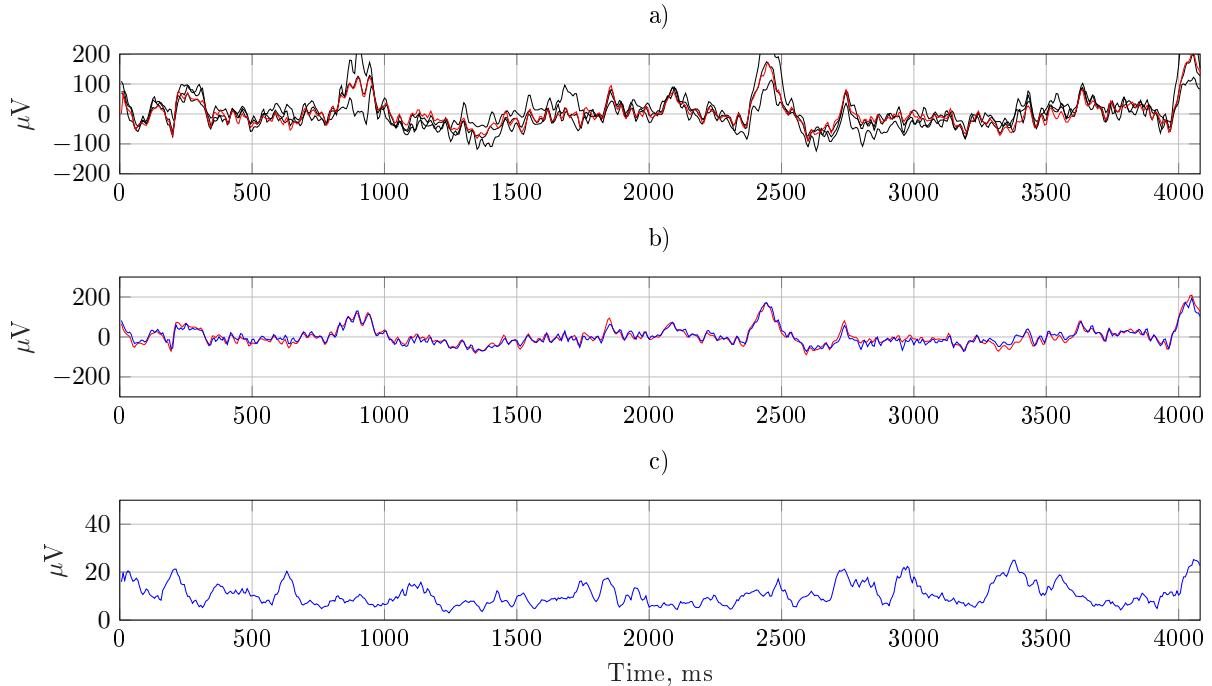


Figure 8. EEG signals and residuals. a) The EEG signal of channel 2 (Red) along with its 4 closest neighbouring channels. b) The predicted (Blue) and observed (Red) signals. c) The residuals with the ocular artefacts occurring at 900 ms and 2300 ms, whose phases are close to the phases at the neighbouring channels. The influence of these artefacts is minimal. The residual is large when the phase difference between the predicted and neighbouring signals is increased as it happened at around 3400 ms.

identification and provide evidence to support Statements 1 and 2 made in section 3. The evidence has been obtained in this study and described in section 6. We also outline possible advantages and limitations of the presented method and finally consider areas of possible applications where the analysis of brain functional connectivity is the research focus.

7.1. Influence of artefacts

EEG signals can be heavily corrupted by cardiac, eye movement and muscle artefacts.^{57,58} Two main types of artefacts are related to eye movement and face muscle activity during the performance of the tasks (tracking of targets on screen). The eye movement artefacts are seen as slow waves (1-2 Hz) with a high amplitude most heavily affecting the frontal electrodes, and the eyebrow muscle, or frontalis, artefacts are observed as 30-40 Hz activity mostly affecting the electrodes AF7/8.⁵⁶ The eye movement artefacts are observed on multiple electrodes without distinguishable phase shifts, and so their influence can

be predicted and minimised within the approach we proposed. In contrast, the frontalis muscle artefacts affect local electrodes and their potentials are more difficult to predict and remove within our approach using the predicted baseline.

On the other hand, the frontalis muscle activity widely varies between individuals in terms of spectral characteristics distributed over electrodes.⁵⁶ Such muscle activity can be associated with a biometric feature of the individual brain print, whose influence on the residuals improves the accuracy of identification.

7.2. Support for Statement 1

Statement 1 has pointed out in section 4.1 that phase delays, which are correlated with the individual brain functional connectivity, are represented by a high-dimensional feature vector which includes the estimates of phase delays between pairs of EEG signals received by a multi-electrode system. In our case the number of electrodes $M = 64$ and therefore the num-

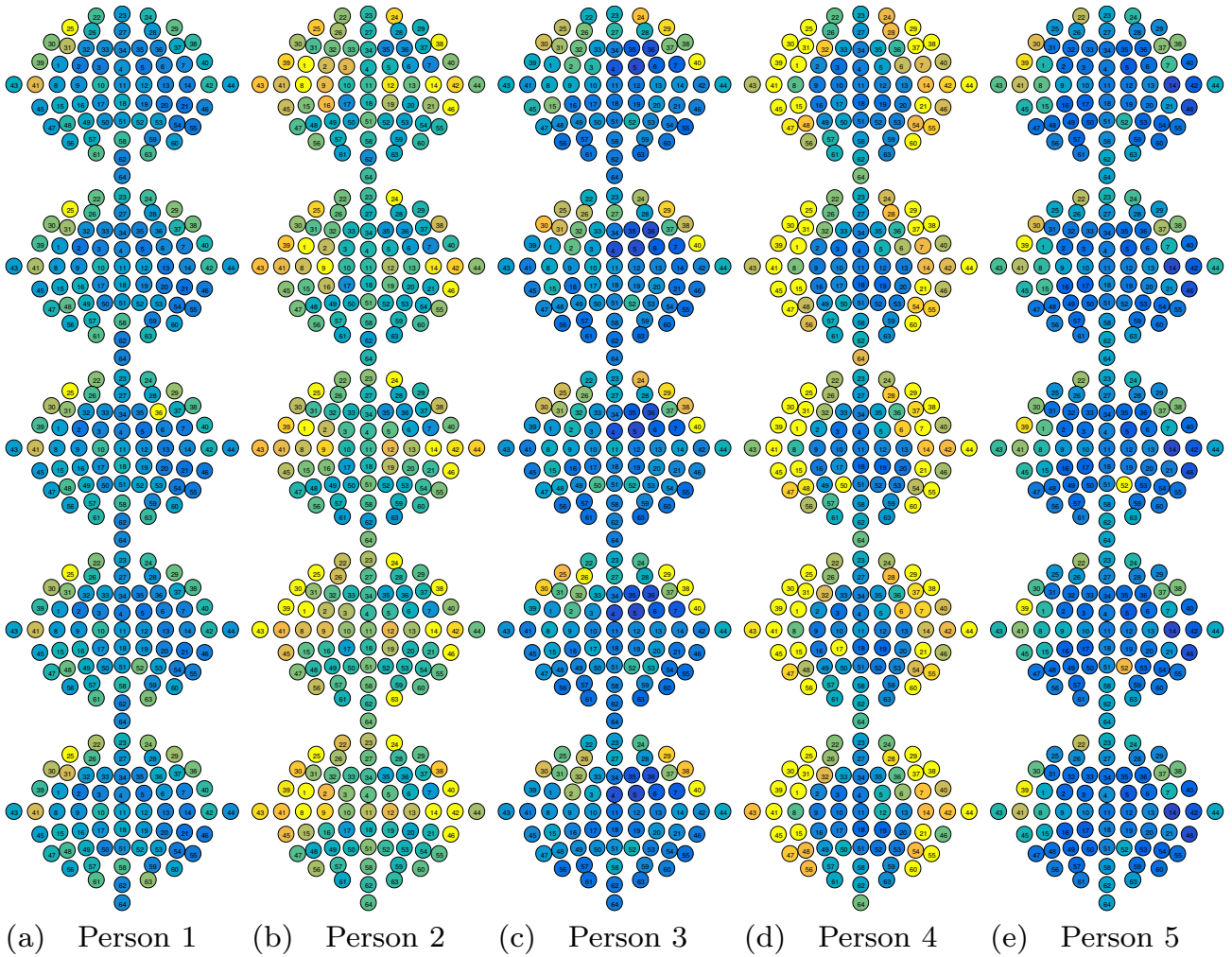


Figure 9. Examples of brain prints discovered for 5 persons from the EEG-MMI data in a sequence of 5 trials. The lowest residual values are in Blue and the highest in Yellow.

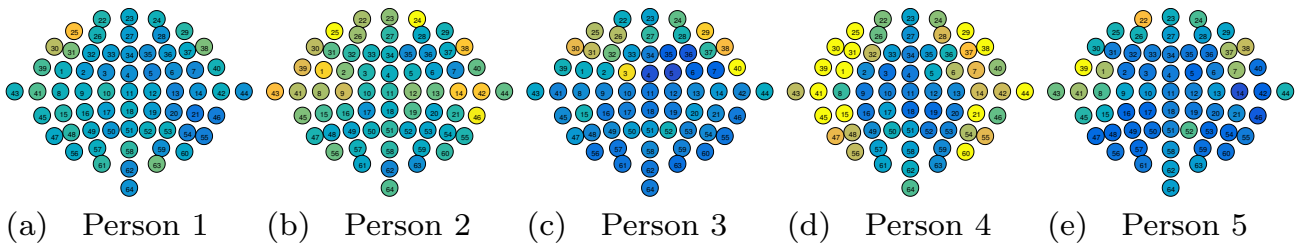


Figure 10. Examples of brain prints averaged over the trials for the above 5 persons.

ber of pairwise comparisons and the dimensionality is 216 for each of frequency component of the EEG. Biometric features obtained in such a way are easily contaminated by noise and artefacts so that some can be irrelevant to the brain print.

The proposed technique, aimed at reducing the

dimensionality of the “phase-based” feature vector, employs the AR-based modelling of the EEG signals received from M electrodes, which delivers a “virtual” baseline required for estimating the individual patterns of brain functional connectivity. As a result, the dimensionality of feature vectors has been signif-

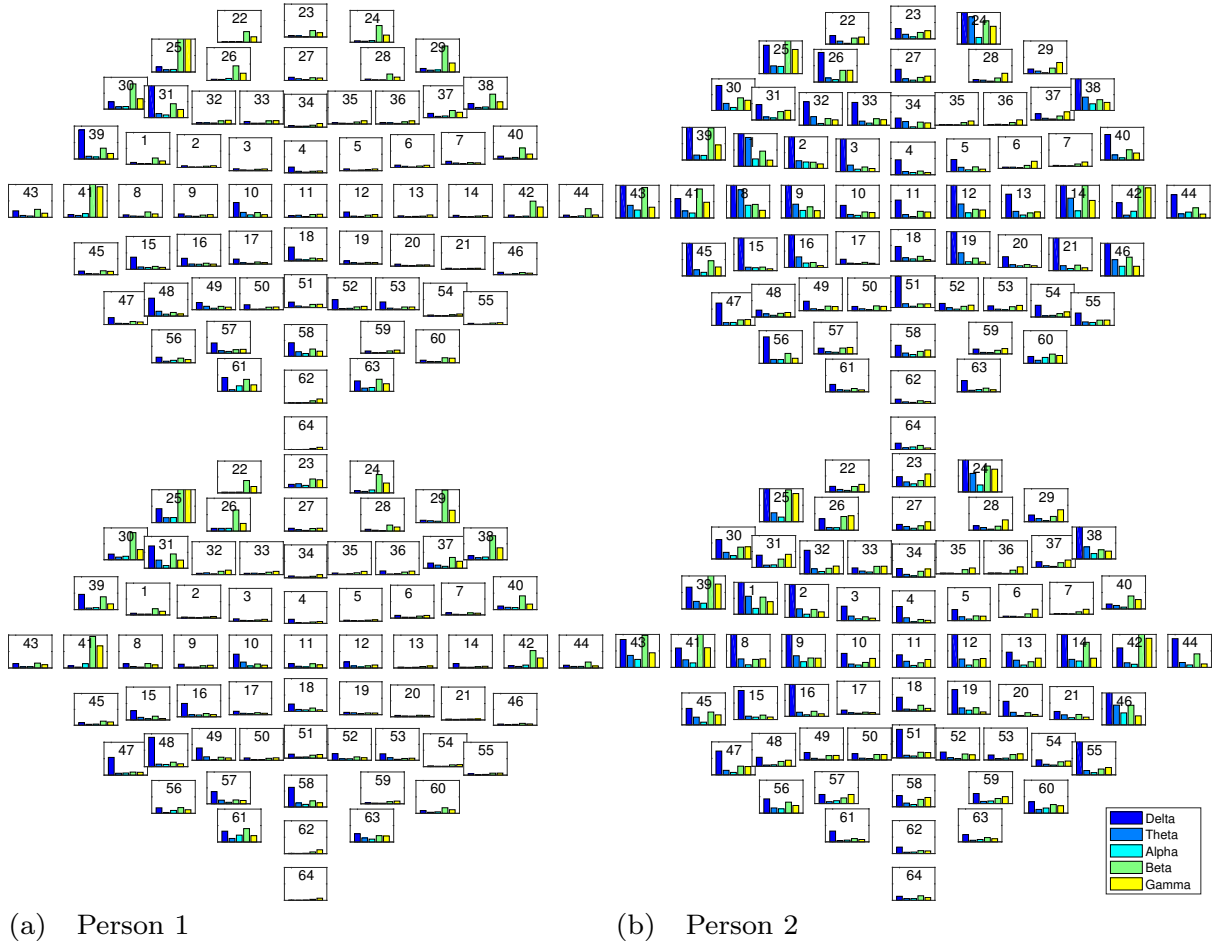


Figure 11. Examples of power spectra distributions of residuals for 2 persons. The distributions are represented in the Delta, Theta, Alpha, Beta, and Gamma bands.

icantly reduced from $2016 \times f$ to 64, where f is the number of frequency components.

Following,⁵⁹ the methodology underlying the proposed method has a fundamental limitation related to the loss of information from signal components caused by fast propagation of neural activity between brain areas under a relatively low sampling rate. In our experiments the EEG-MMI data have a sampling rate of 160 Hz, and so the event delays that are shorter than 6.25 ms will not be captured by the method. Therefore the components of such activity cannot be distinguished from the volume conduction. However in our experiments the loss of these components has a minor effect and the identification accuracy, shown in Table 1, has not been affected.

The above limitation is inherent in Hjorth's source derivation method outlined in section 3.2.

Both methods aim to decorrelate the signals, however the differences between the source derivation and proposed methods can be seen from Eq. 1 and Eq. 14, and the corresponding block diagrams 2 and 3, as follows. The source derivation method in its advanced adaptive form⁵² requires the inverse operator to calculate the sensor matrix \mathbf{B} in Eq. 3, which can produce unstable derivation of the source. The proposed method employs AR modelling to orthogonalise signals of brain activities received by EEG electrodes from different brain areas. In this regard, it is important to note that AR modelling has been shown efficient for discovering directions of information flows in brain structures.⁶⁰ Neither the source derivation method⁵⁰ nor its extensions^{51,52} consider this important component of the brain functional activity. We can also assume that both methods use

the same 5-electrode placement without delayed signals so that Eq. 14 has $M = 5$ and $P = 0$. Then the source derivation method⁵⁰ produces outcomes determined by Eq. 1 for the predefined unit coefficients. In contrast, the proposed method produces the outcomes determined by Eq. 14 with the coefficients β_1, \dots, β_M , which have been fitted to the signal of a central electrode i . Taking into account the above differences, we conclude that the proposed method is more advanced for modelling the brain connectivity.

The AR-based models have efficiently suppressed the noise and artefacts we discussed in section 7.1. Fig. 8 shows that the artefacts, which are caused by eye movements at around 900 and 2,300 ms, do not affect the virtual baseline (the Blue line on the middle plot) and so their residuals (on the bottom plot) do not exceed the average value.

Based on the experimental evidence we conclude that the phases mapped in this way onto the modelling residuals, which are used as biometric features, are capable of making important contributions to the identification.

7.3. Support for Statement 2

The Statement 2 has pointed out in section 4.2 that the enrolment data, which are represented by a matrix \mathbf{X} , can be underdetermined, and columns of the matrix can be correlated. These conditions make finding of a numerically stable estimate of the coefficient vector $\hat{\mathbf{w}}$ problematic.

The GMDH-based technique has been stated to be capable of overcoming this problem by using the regularisation criterion. This criterion is based on splitting the training data into 2 disjoint subsets, namely training \mathcal{A} and validation \mathcal{B} , that has been shown efficient to find reliable estimates of $\hat{\mathbf{w}}$. In our experiments on the benchmark data the desired regularisation is achieved by evaluating the residual Δ of the model $f(\cdot; \alpha_{\mathcal{A}}, \mathcal{X})$ on the entire enrolment data that include the validation subset \mathcal{B} .

Besides the above problems, the AR modelling requires determination of a structural vector \mathcal{X} which describes the interactions between channels. In our experiments the proposed GMDH algorithm has found the structure vector with the largest contribution made from 4 electrodes neighbouring a central

electrode. It is interesting that this finding matches the results⁵⁰ reported by Hjorth.

Under these conditions the reliable modelling of the multichannel interregional interaction, which includes the contribution of components of AR model can be achieved by Eq. 14 with the extended parameter vectors $(\hat{\beta}, \hat{\alpha}, \hat{\mathcal{X}})$. These vectors can be learnt from the given enrolment data by using the proposed GMDH algorithm. In our experiments we found that the proposed algorithm is capable of delivering a reliable solution to the minimization problem (Eq. 16).

In this relation the modelling of the multichannel interaction can be seen as an integration over brain regions:

$$\begin{aligned} x(t) &= \sum_{i=1}^M \beta_i \sum_{j=1}^{P_i} \alpha_{ij} x(t-j) + r(t) \\ &= \sum_{i=1}^{M_p} \alpha_i x(t-j) + r(t), \end{aligned}$$

where M_p is the number of weight coefficients included in the vector α .

In our study, given $M = 64$ electrodes and $P = 5$ lags, the maximum M_p is 320, which is sufficient to achieve the accurate extraction of the brain print features from the multi-channel EEG data presented for the enrolment.

A limitation of our study is that the experiments were performed on EEG data recorded on a single session for each individual. This limitation, however, is inherent to all studies on EEG biometrics discussed in section 2 which were performed on the EEG-MMI data, the largest publicly available benchmark.

7.4. Possible application areas and future work

Biometric methods have shown the potential to improve security and convenience in many areas, including health care. Recent work has shown that EEG signals are distinctive enough for each individual person to be used for biometric applications. In health care, patient misidentification is a real problem. In the UK, more than 24,000 cases of patients were reported mismatched to their care between 2006 and 2007. The NHS lists patient misidentification as serious patient safety incidents that should not occur.⁶¹

EEG is an important tool for the assessment of brain development in newborns.^{62–65} It has been found that coherent EEG activity during sleep may provide unique insight into maturation processes of brain functional connectivity. Assessments have confirmed that sleep EEG coherence increases across development.⁶⁶

Brain connectivity in patients with psychiatric disorders has been investigated using EEG techniques enabling the exploration of oscillatory brain dynamics. The results from the EEG studies clarify the role of neuronal network synchronisation as a potential biomarker of bipolar,⁶⁷ depressive^{68,69} and autism⁵⁸ disorders.

Attention deficit disorders (ADD) that interfere with successful lives cannot be reliably diagnosed with MRI. This is because ADD is not only an anatomical pattern of brain abnormality, but also reflects a unique response to different pathology causes. ADD-specific brain patterns may be better detected by measuring EEG coherence and cortical connectivity.^{70–72}

Brain functional connectivity being represented by statistical dependencies between the dynamics of recorded EEG signals is used to analyse the dynamical interactions of brain regions. The functional connectivity patterns obtained from the scalp EEG recordings reveal information about the dynamics of the epileptic brain and can be used to predict upcoming seizures.⁷³

Alzheimer’s disease (AD), which is the common type of neurodegenerative disorder, targets cortical neuronal networks related to cognitive functions. AD is partially reflected by the abnormal mechanisms of cortical neural synchronisation that generate resting state EEG rhythms. The cortical neural synchronisation can be therefore represented by EEG power densities representing the brain functional connectivity.⁷⁴

Neural activity during movement tasks carries important information for Brain-Computer Interface (BCI). It has been shown that the functional connectivity can be measured to discriminate hand movements and resting state conditions. The analysis of the functional connectivity could be an efficient alternative to conventional BCI techniques.⁷⁵

8. Conclusion

Methods of EEG-based person identification and extraction of biometric features have been analysed in terms of reliability and accuracy. The analysis has revealed that EEG signals represent the brain activity which is mainly influenced by the volume conduction and functional connectivity. The first component is not relevant to the brain print because of trivial correlations between electrodes, while the second component determined by the “connectotype” of a person. In the presence of the volume conduction, this component can be estimated after orthogonalisation of the multichannel EEG data. The results are represented as a high-dimensional matrix of interchannel interactions. The biometric features which are relevant to the brain connectivity can be extracted from the residuals (Eq. 8). The model parameters are the solution of the minimisation problem (Eq. 13) that is delivered by the orthogonalisation.

However the conventional methods of orthogonalisation of high-dimensional EEG data in the presence of electrode noise and muscle artefacts cannot provide reliable solutions for the minimisation problem (Eq. 6), which is required to correlate the residuals with the connectotype of a person during a reasonably short test. We found that the dimensionality can be significantly reduced if the baselines which are required for estimating the residuals can be reliably modelled by using electrodes neighbouring a central electrode.

These findings have led to our new approach to learning of the required models by using a GMDH algorithm, which is capable of discovering models that provide the maximal generalisation ability in the case of multidimensional EEG data. In our experiments on the EEG-MMI data benchmark with 109 participants, the proposed method has shown 100% identification accuracy using short 4-sec tests, thus outperforming existing biometric approaches.

The above discussion and findings allow us to conclude that the proposed method is a promising approach to achieving reliable results of EEG-based person identification in many realistic cases, and this will be further explored in a wider scope of relevant biometric applications and tasks.

Acknowledgments

The authors are grateful to the anonymous reviewers for their constructive comments. The research has been partly supported by the UK Leverhulme Trust, Grant F/00 811/A.

Bibliography

1. R. Palaniappan, Two-stage biometric authentication method using thought activity brain waves, *International Journal of Neural Systems* **18**(01) (2008) 59–66.
2. H. P. Da Silva, A. Fred and K. Ricanek, Harnessing the power of biosignals, *Computer* **47**(3) (2014) 74–77.
3. S. Marcel and J. D. R. Millan, Person authentication using brainwaves (EEG) and maximum a posteriori model adaptation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **29**(4) (2007) 743–752.
4. E. Maiorana, D. L. Rocca and P. Campisi, Eigenbrains and eigentensorbrains: Parsimonious bases for EEG biometrics, *Neurocomputing* **171** (2016) 638–648.
5. H. Stassen, Computerized recognition of persons by EEG spectral patterns, *Electroencephalography and Clinical Neurophysiology* **49**(1–2) (1980) 190–94.
6. R. B. Paranjape, J. Mahovsky, L. Benedicenti and Z. Koles, The electroencephalogram as a biometric, *Canadian Conference on Electrical and Computer Engineering*, **2**, (Toronto, Ontario, Canada, 2001), pp. 1363–1366.
7. M. DelPozo-Banos, C. M. Travieso, C. T. Weidemann and J. B. Alonso, EEG biometric identification: a thorough exploration of the time-frequency domain, *Journal of neural engineering* **12**(5) (2015) p. 056019.
8. O. Miranda-Dominguez, B. D. Mills, S. D. Carpenter, K. A. Grant, C. D. Kroenke, J. T. Nigg and D. A. Fair, Connectotyping: Model based fingerprinting of the functional connectome, *PLOS ONE* **9**(11) (2014).
9. M. Fraschini, A. Hillebrand, M. Demuru, L. Didaci and G. L. Marcialis, An EEG-based biometric system using eigenvector centrality in resting state brain networks, *IEEE Signal Processing Letters* **22** (June 2015) 666–670.
10. A. Crobe, M. Demuru, L. Didaci, G. L. Marcialis and M. Fraschini, Minimum spanning tree and k-core decomposition as measure of subject-specific EEG traits, *Biomedical Physics & Engineering Express* **2**(1) (2016).
11. D. J. A. Smit, C. J. Stam, D. Posthuma, D. I. Boomsma and E. J. C. De Geus, Heritability of "small-world" networks in the brain: A graph theoretical analysis of resting-state EEG functional connectivity, *Human Brain Mapping* **29**(12) (2008) 1368–1378.
12. D. J. A. Smit, M. Boersma, C. E. M. van Beijsterveldt, D. Posthuma, D. I. Boomsma, C. J. Stam and E. J. C. de Geus, Endophenotypes in a dynamically connected brain, *Behavior Genetics* **40**(2) (2010) 167–177.
13. M. Ahmadlou and H. Adeli, Visibility graph similarity: A new measure of generalized synchronization in coupled dynamic systems, *Physica D: Nonlinear Phenomena* **241**(4) (2012) 326–332.
14. A. Riera, A. Soria-Frisch, M. Caparrini, C. Grau and G. Ruffini, Unobtrusive biometric system based on electroencephalogram analysis, *EURASIP Journal on Advances in Signal Processing* **2008**(1) (2007) p. 143728.
15. D. L. Rocca, P. Campisi, B. Vegso, P. Cserti, G. Kozmann, F. Babiloni and F. D. V. Fallani, Human brain distinctiveness based on EEG spectral coherence connectivity, *IEEE Transactions on Biomedical Engineering* **61** (Sept 2014) 2406–2412.
16. G. Alarcón, C. D. Binnie, S. J. J. García, M. M. C. Martin, T. J. L. Fernandez, C. E. Polkey and C. N. Guy, Mechanisms involved in the propagation of interictal epileptiform discharges in partial epilepsy, *Clinical Neurophysiology: from receptors to Perception*, eds. G. Comi, C. Lucking, J. Kimura and P. Rossini, *Electroencephalography and Clinical Neurophysiology Supplements* **50** (Elsevier, 1999), pp. 259–278.
17. G. Nolte, O. Bai, L. Wheaton, Z. Mari, S. Vorbach and M. Hallett, Identifying true brain interaction from EEG data using the imaginary part of coherency, *Clinical Neurophysiology* **115**(10) (2004) 2292–2307.
18. C. J. Stam, G. Nolte and A. Daffertshofer, Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources, *Human Brain Mapping* **28**(11) (2007) 1178–1193.
19. M. Brookes, M. Woolrich and G. Barnes, Measuring functional connectivity in MEG: A multivariate approach insensitive to linear source leakage, *NeuroImage* **63**(2) (2012) 910–920.
20. J. F. Hipp, D. J. Hawellek, M. Corbetta, M. Siegel and A. K. Engel, Large-scale cortical correlation structure of spontaneous oscillatory activity, *Nature Neuroscience* **15**(6) (2012) 884–890.
21. C. Liu, B. Abu-Jamous, E. Brattico and A. K. Nandi, Towards tunable consensus clustering for studying functional brain connectivity during affective processing, *International Journal of Neural Systems* **27**(02) (2017) p. 1650042.
22. H. R. Madala and A. G. Ivakhnenko, *Inductive Learning Algorithms for Complex Systems Modeling* (CRC Press, Boca Raton, 1994).
23. D. J. Hand, Pattern detection and discovery, *Pattern Detection and Discovery: ESF Exploratory Work-*

- shop, eds. D. J. Hand, N. M. Adams and R. J. Bolton (Springer, 2002).
24. T. Hastie, R. Tibshirani and J. Friedman, *The Elements of Statistical Learning* (Springer-Verlag New York, 2009).
 25. L. Breiman, J. Friedman, R. Olshen and C. Stone, *Classification and Regression Trees* (Chapman and Hall, 1984).
 26. A. Ivakhnenko, Polynomial theory of complex systems, *IEEE Transactions on Systems, Man and Cybernetics SMC-1(4)* (1997) 364–378.
 27. S. J. Farlow, *Self-Organizing Methods in Modelling: GMDH Type Algorithms* (Marcel Decker, 1984).
 28. J. Schmidhuber, Deep learning in neural networks: An overview, *Neural Networks* **61** (2015) 85–117.
 29. Y. LeCun, Y. Bengio and G. Hinton, Deep learning, *Nature* **521** (May 2015) 436–444.
 30. F. C. Morabito, M. Campolo, N. Mammone, M. Versaci, S. Franceschetti, F. Tagliavini, V. Sofia, D. Fatuzzo, A. Gambardella, A. Labate, L. Mumoli, G. G. Tripodi, S. Gasparini, V. Cianci, C. Sueri, E. Ferlazzo and U. Aguglia, Deep learning representation from electroencephalography of early-stage Creutzfeldt-Jakob disease and features for differentiation from rapidly progressive dementia, *International Journal of Neural Systems* **27**(02) (2017) p. 1650039.
 31. V. Schetinin and J. Schult, Learning polynomial networks for classification of clinical electroencephalograms, *Soft Computing* **10** (Feb 2006) 397–403.
 32. V. Schetinin and J. Schult, A neural-network technique to learn concepts from electroencephalograms, *Theory in Biosciences* **124** (Aug 2005) 41–53.
 33. T. Gasser, P. Bächer and H. Steinberg, Test-retest reliability of spectral parameters of the EEG, *Electroencephalography and Clinical Neurophysiology* **60**(4) (1985) 312–319.
 34. M. Salinsky, B. Oken and L. Morehead, Test-retest reliability in EEG frequency analysis, *Electroencephalography and Clinical Neurophysiology* **79**(5) (1991) 382–392.
 35. R. Palaniappan and D. P. Mandic, Biometrics from brain electrical activity: A machine learning approach, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **29** (April 2007) 738–742.
 36. G. Bajwa and R. Dantu, Neurokey: Towards a new paradigm of cancelable biometrics-based key generation using electroencephalograms, *Computers & Security* **62** (2016) 95–113.
 37. G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer and J. R. Wolpaw, BCI2000: a general-purpose brain-computer interface (BCI) system, *IEEE Transactions on Biomedical Engineering* **51** (June 2004) 1034–1043.
 38. M. del Pozo Banos, C. M. Travieso, J. B. Alonso and A. John, Evidence of a task-independent neural signature in the spectral shape of the electroencephalogram, *International Journal of Neural Systems* **28**(1) (2018) p. 1750035.
 39. D. Rodrigues, G. F. Silva, J. P. Papa, A. N. Marana and X.-S. Yang, EEG-based person identification through binary flower pollination algorithm, *Expert Systems with Applications* **62** (2016) 81–90.
 40. S. Yang, F. Deravi and S. Hoque, Task sensitivity in EEG biometric recognition, *Pattern Analysis and Applications* **1** (2016) 1–13.
 41. K. J. Friston, Functional and effective connectivity in neuroimaging: A synthesis, *Human Brain Mapping* **2**(1-2) (1994) 56–78.
 42. M. Ahmadi, A. Adeli, R. Bajo and H. Adeli, Complexity of functional connectivity networks in mild cognitive impairment subjects during a working memory task, *Clinical Neurophysiology* **125**(4) (2014) 694–702.
 43. E. S. Finn, X. Shen, D. Scheinost, M. D. Rosenberg, J. Huang, M. M. Chun, X. Papademetris and R. T. Constable, Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity, *Nature Neuroscience* **18** (October 2015) 1664–1671.
 44. M. Leite, A. Leal and P. Figueiredo, Transfer function between EEG and BOLD signals of epileptic activity, *Frontiers in Neurology* **4** (2013) p. 1.
 45. V. Menon, W. Freeman, B. Cuttillo, J. Desmond, M. Ward, S. Bressler, K. Laxer, N. Barbaro and A. Gevins, Spatio-temporal correlations in human gamma band electrocorticograms, *Electroencephalography and Clinical Neurophysiology* **98**(2) (1996) 89–102.
 46. P. L. Nunez, R. Srinivasan, A. F. Westdorp, R. S. Wijesinghe, D. M. Tucker, R. B. Silberstein and P. J. Cadusch, EEG coherence: I: statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales, *Electroencephalography and Clinical Neurophysiology* **103**(5) (1997) 499–515.
 47. J.-P. Lachaux, E. Rodriguez, J. Martinerie and F. J. Varela, Measuring phase synchrony in brain signals, *Human Brain Mapping* **8**(4) (1999) 194–208.
 48. P. L. Nunez, *Neocortical Dynamics and Human EEG Rhythms* (Oxford University Press, 1995).
 49. R. W. Thatcher, Coherence, phase differences, phase shift, and phase lock in EEG/ERP analyses, *Developmental Neuropsychology* **37**(6) (2012) 476–496.
 50. B. Hjorth, An on-line transformation of EEG scalp potentials into orthogonal source derivations, *Electroencephalography and Clinical Neurophysiology* **39**(5) (1975) 526–530.
 51. B. Hjorth, Physical aspects of EEG data as a basis for topographic mapping, *Topographic Mapping of Brain Electrical Activity*, ed. F. Duffy (Butterworth-Heinemann, 2013), pp. 175–194.
 52. B. Hjorth, An adaptive EEG derivation technique, *Clinical Neurophysiology* **54** (1982) 654–661.
 53. A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus,

- G. B. Moody, C.-K. Peng and H. E. Stanley, PhysioBank, PhysioToolkit, and PhysioNet, *Circulation* **101**(23) (2000) e215–e220.
54. J. Fürnkranz, Round robin classification, *Journal of Machine Learning Research* **2** (March 2002) 721–747.
 55. J. Uglov, L. Jakaite, V. Schetinin and C. Maple, Comparing robustness of pairwise and multi-class neural-network systems for face recognition, *EURASIP Journal on Advances in Signal Processing* **2008** (Dec 2007).
 56. I. Goncharova, D. McFarland, T. Vaughan and J. Wolpaw, EMG contamination of EEG: spectral and topographical characteristics, *Clinical Neurophysiology* **114**(9) (2003) 1580–1593.
 57. V. Schetinin and C. Maple, A Bayesian model averaging methodology for detecting EEG artifacts, *2007 15th International Conference on Digital Signal Processing*, (Cardiff, 2007), pp. 499–502.
 58. M. Ahmadlou, H. Adeli and A. Adeli, Improved visibility graph fractality with application for the diagnosis of autism spectrum disorder, *Physica A: Statistical Mechanics and its Applications* **391**(20) (2012) 4720–4726.
 59. G. Alarcón, J. J. Garcia Seoane, C. D. Binnie, M. C. Martin Miguel, J. Juler, C. E. Polkey, R. D. Elwes and J. M. Ortiz Blasco, Origin and propagation of interictal discharges in the acute electrocorticogram. implications for pathophysiology and surgical treatment of temporal lobe epilepsy, *Brain* **120**(12) (1997) 2259–2282.
 60. A. Korzeniewska, M. Mańczak, M. Kamiński, K. J. Blinowska and S. Kasicki, Determination of information flow direction among brain structures by a modified directed transfer function (dDTF) method, *Journal of Neuroscience Methods* **125**(1–2) (2003) 195–207.
 61. A. Rinaldi, Biometrics’ new identity: Measuring more physical and biological traits, *EMBO reports* **17**(1) (2016) 22–26.
 62. L. Jakaite, V. Schetinin and J. Schult, Feature extraction from electroencephalograms for Bayesian assessment of newborn brain maturity, *24th International Symposium on Computer-Based Medical Systems (CBMS)*, (Bristol, 2011), pp. 1–6.
 63. V. Schetinin and L. Jakaite, Classification of newborn EEG maturity with Bayesian averaging over decision trees, *Expert Systems with Applications* **39**(10) (2012) 9340–9347.
 64. L. Jakaite, V. Schetinin and C. Maple, Bayesian assessment of newborn brain maturity from two-channel sleep electroencephalograms, *Computational and Mathematical Methods in Medicine* **2012** (2012) 1–7.
 65. V. Schetinin and L. Jakaite, Extraction of features from sleep EEG for Bayesian assessment of brain development, *PLOS ONE* **12** (03 2017) 1–13.
 66. S. Kurth, P. Achermann, T. Rusterholz and M. LeBourgeois, Development of brain EEG connectivity across early childhood: Does sleep play a role?, *Brain Sciences* **3**(4) (2013) 1445–1460.
 67. E. Maggioni, A. Bianchi, A. Altamura, J. C. Soares and P. Brambilla, The putative role of neuronal network synchronization as a potential biomarker for bipolar disorder: A review of EEG studies, *Journal of Affective Disorders* **212** (2017) 167–170.
 68. M. Ahmadlou, H. Adeli and A. Adeli, Fractality analysis of frontal brain in major depressive disorder, *International Journal of Psychophysiology* **85**(2) (2012) 206–211.
 69. M. Ahmadlou, H. Adeli and A. Adeli, Spatiotemporal analysis of relative convergence of EEGs reveals differences between brain dynamics of depressive women and men, *Clinical EEG and Neuroscience* **44**(3) (2013) 175–181.
 70. F. H. Duffy, A. Shankardass, G. B. McAnulty and H. Als, A unique pattern of cortical connectivity characterizes patients with attention deficit disorders: a large electroencephalographic coherence study, *BMC Medicine* **15** (Mar 2017).
 71. M. Ahmadlou and H. Adeli, Functional community analysis of brain: A new approach for EEG-based investigation of the brain pathology, *NeuroImage* **58**(2) (2011) 401–408.
 72. M. Ahmadlou, H. Adeli and A. Adeli, Graph theoretical analysis of organization of functional brain networks in ADHD, *Clinical EEG and Neuroscience* **43**(1) (2012) 5–13.
 73. P. van Mierlo, M. Papadopoulou, E. Carrette, P. Boon, S. Vandenberghe, K. Vonck and D. Marinazzo, Functional brain connectivity from EEG in epilepsy: Seizure prediction and epileptogenic focus localization, *Progress in Neurobiology* **121** (2014) 19–35.
 74. C. Babiloni, R. Lizio, N. Marzano, P. Capotosto, A. Soricelli, A. I. Triggiani, S. Cordone, L. Gesualdo and C. D. Percio, Brain neural synchronization and functional coupling in Alzheimer’s disease as revealed by resting state EEG rhythms, *International Journal of Psychophysiology* **103** (2016) 88–102.
 75. M. Demuru, F. Fara and M. Fraschini, Brain network analysis of EEG functional connectivity during imagery hand movements, *Journal of Integrative Neuroscience* **12**(04) (2013) 441–447.