

DOI: 10.5604/20830157.1130196

## DWUMODUŁOWY SYSTEM DO PRZETWARZANIA DANYCH EEG Z WYKORZYSTANIEM ANALIZY CZYNNIKOWEJ I PSEUDOINWERSJI MOORE-PENROSE

Szczepan Paszkiel

Politechnika Opolska, Wydział Elektrotechniki, Automatyki i Informatyki, Instytut Automatyki i Informatyki

**Streszczenie.** W artykule opisano koncepcję otrzymywania tzw. sygnału wyjściowego na potrzeby między innymi zastosowań do procesów sterowania. W tym celu zaproponowano budowę dwumodułowego systemu do przetwarzania i analizy danych elektroencefalograficznych w skład którego wchodzi analiza czynnikowa oraz pseudoinwersja Moore-Penrose. W artykule scharakteryzowano także problem dużej interferencyjności źródeł sygnałów EEG, co ma negatywny wpływ na finalny proces otrzymywania wyjściowego sygnału z zastosowaniem w automatyce, czy robotyce. Implikuje to tym samym konieczność właściwej i poprawnej identyfikacji źródeł sygnału w mózgu człowieka.

**Słowa kluczowe:** system dwumodułowy, pseudoinwersja Moore-Penrose, analiza czynnikowa, dane EEG

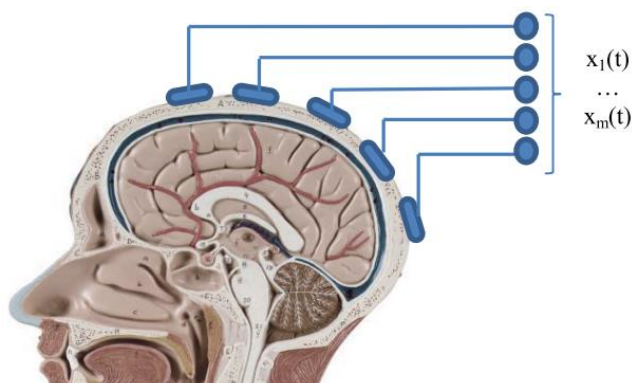
### TWO-MODULAR SYSTEM FOR PROCESSING EEG DATA USING FACTOR ANALYSIS AND MOORE-PENROSE PSEUDOINVERSION

**Abstract.** This paper describes the concept of obtaining the so-called. the output signal for the purpose, inter alia, the control processes carried out. To this end, proposed the construction of two-modular system for processing and analysis of electrophysiological data on the composition, which includes factor analysis and pseudoinversion Moore Penrose. In the article the problem of high interference sources of EEG signals, which has a negative impact on the process of obtaining the final output using automation or robotics. This implies also the problem of proper and correct identification of sources in the human brain.

**Keywords:** two-modular system, Moore-Penrose pseudoinversion, factor analysis, EEG data

#### Wstęp

Mózg człowieka jest złożoną jednostką obliczeniową. Składa się z miliardów neuronów, poprzez których wzajemne korelacje generowane są potencjały mózgowie, możliwe do odczytu za pomocą urządzeń jakimi są między innymi elektroencefalografy. Jest to jedna z wielu metod badania aktywności pracy mózgu, której cechą jest nieinwazyjność, a także brak negatywnego wpływu na zdrowie człowieka przy dużej możliwości powtarzalności pomiarów. W badaniu elektroencefalograficznym na powierzchni głowy umieszcza się elektrody, które wykorzystywane są do przeprowadzenia procesu akwizycji sygnału biologicznego. Na rysunku 1 przedstawiono schematycznie ułożenie elektrod. Zarejestrowane przebiegi oznaczone  $x_1(t) \dots x_m(t)$  identyfikują sygnały obserwowane za pomocą czujników pomiarowych – elektrod [3].



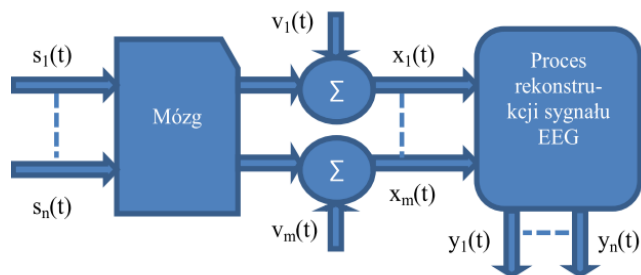
Rys. 1. Schematyczne umiejscowienie elektrod na głowie osoby badanej

Nieustanny rozwój technologii mózg-komputer implikuje konieczność rozbudowy istniejących dotychczas metod analizy sygnałów elektroencefalograficznych i stosowania nowych rozwiązań matematycznych, w tym także budowę systemów modułowych łączących funkcjonalności kilku metod. Tym sposobem rozwiązania mające na celu bezdotykowe przeszukiwanie Internetu, sterowanie pojazdami mobilnymi, komputerami przemysłowymi, a także udoskonalone metody uwierzytelniania nabierają nowej jakości i dokładności w ich realizacji. Jak wynika z badań zapotrzebowanie na nowoczesne

i innowacyjne metody komunikacji człowiek-komputer, człowiek-maszyna jest coraz to większe, szczególnie, że u niektórych osób (dotkniętych między innymi chorobą Charcota, chorobą Lou Gehriga etc.) może być to jedyna, dotychczas jeszcze nie spopularyzowana na wielką skalę metoda nawiązania kontaktu ze światem zewnętrznym.

#### 1. Identyfikacja źródeł sygnału EEG

Złożoność ludzkiego mózgu jest bardzo duża, o czym świadczyć może między innymi ogromna liczba synaps, jakie w nim występują. Jest ona rzędu  $10^{14}$ , co pozwala na wykonywanie w przeliczeniu  $10^{15}$  operacji na sekundę [2]. Sygnały elektroencefalograficzne, mające swoje źródło w mózgu podlegają interferencyjności. Ich nakładanie się na siebie, znacząco utrudnia pomiar sygnału na powierzchni czaszki. Tym samym jeszcze trudniej jest dokonać procesu przybliżonej lokalizacji źródła w strukturze mózgu. Niestety sygnał EEG w przypadku metody nieinwazyjnej, na potrzeby której opisano w niniejszym artykule system dwumodułowy, jest dodatkowo filtrowany przez warstwy: kości, płynu mózgowo-rdzeniowego oraz opon mózgowych [9]. Typowymi przykładami filtru reaktancyjnego są w tym przypadku skóra oraz tkanki podskórne jakie znajdują się na głowie osoby poddawanej badaniu elektroencefalograficznemu. Elementem czynnym takich filtrów jest struktura hydroksyapatytu oraz elektrolitów wielokrotnie przegradzanych warstwami hydrofilno-hydrofobowymi [10].



Rys. 2. Przedstawienie procesu rekonstrukcji sygnału źródłowego

Na rysunku 2 przedstawiono w sposób ogólny poszczególne etapy powstawania biologicznego wyjściowego sygnału źródłowego. Niezliczona liczba źródeł sygnałów znajdujących się

w mózgu –  $s(t)$  w połączeniu z szeregiem zakłóceń –  $v(t)$ , implikuje powstanie sygnałów zarejestrowanych za pomocą czujników – elektrod oznaczonych jako  $x(t)$ . Następnie na potrzeby prowadzonych analiz realizowana jest odpowiednio rekonstrukcja sygnałów na stacji roboczej poprzez zastosowanie adekwatnych algorytmów wchodzących w skład proponowanego w niniejszym artykule dwumodułowego systemu analizy danych EEG z wykorzystaniem analizy czynnikowej oraz pseudoinwersji Moore-Penrosa [7]. Tym sposobem na wyjściu uzyskuje się sygnał  $y(t)$ ,  $x(t)=[x_1(t), \dots, x_m(t)]^T$  – to sygnały zarejestrowane za pomocą czujników,  $s(t)=[s_1(t), \dots, s_n(t)]^T$  – sygnały pierwotne EEG – źródła sygnałów w mózgu,  $y(t)=[y_1(t), \dots, y_n(t)]^T$  – sygnały wyjściowe, uzyskane po rekonstrukcji.  $n$  – liczba źródeł,  $m$  – liczba obserwacji prowadzonych w ramach eksperymentu pomiarowego.

## 2. Zastosowanie analizy czynnikowej na potrzeby rekonstrukcji źródłowych sygnałów EEG

Faktoryzacja to proces w którym dla danego obiektu możliwe są do znalezienia takie obiekty, których iloczyn jest równy wyżej wymienionemu [5, 6]. Analiza czynnikowa należy z kolei do grupy metod statystycznych [1]. Najczęściej wykorzystuje się ją w teorii decyzji, marketingu i zarządzaniu. Jak dowiodły jednak przeprowadzone badania analizę czynnikową z powodzeniem można zastosować do przetwarzania sygnałów biomedycznych, takich jak sygnał elektroencefalograficzny. Główny cel analizy czynnikowej to redukcja liczby zmiennych. Jest to możliwe poprzez założenie, że pewna grupa zmiennych losowych przedstawia zmienność tych samych czynników. Pozwala to na stwierdzenie, że zmienne losowe są od siebie niezależne. Jedną z metod wyżej wymienionej analizy jest analiza składowych głównych PCA (ang. *Principal Component Analysis*). Główny cel metody to rotacja układu współrzędnych w takim stopniu, aby uzyskać maksymalizację wariancji pierwszej współrzędnej, a następnie kolejnych [15, 19].

Posiadając przykładowy zbiór danych wejściowych w postaci sygnału EEG, który traktujemy jako  $N$  obserwacji, z których to każda obejmuje określoną liczbę zmiennych  $M$ , można przyjąć że  $N$  określa punkty w przestrzeni  $M$ -wymiarowej. Tym sposobem uzyskujemy nową przestrzeń obserwacji, którą określają początkowe czynniki. Metoda ta w przypadku analizy sygnału elektroencefalograficznego pozwala na zmniejszenie ilości informacji zawartych w sygnale, poprzez odrzucenie pewnych składowych, zawierających artefakty zakłócające [11]. Podczas zastosowania powyższej metody możliwe jest przedstawienie zbioru wejściowego w postaci macierzy kowariancji lub korelacji.

Na rysunku 3 przedstawiono sposób konstruowania wektorów  $x(k)$ , które to reprezentują macierz faktoryzacji  $X$ . Macierz faktoryzacji tworzona jest poprzez iloczyn macierzy mieszającej  $A$  z sygnałami źródłowymi  $S$  w sumie z macierzą reprezentującą błędy i zakłócenia  $V$ . Opisuje to równanie (1):

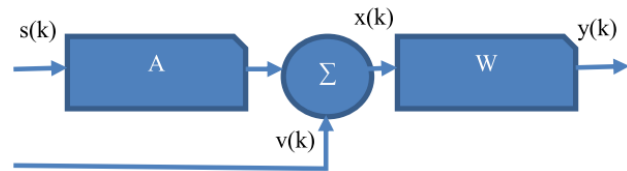
$$X = AS + V \quad (1)$$

gdzie:  $X=[x(1), \dots, x(N)]$  – macierz faktoryzacji reprezentowana przez wektory  $x(k)$ ,  $k=1, 2, \dots, N$ ;  $k=t$ ;  $S=[s(1), \dots, s(N)]$  – ukryte komponenty reprezentujące sygnały źródłowe;  $N$  – liczba dostępnych próbek,  $A$  – macierz mieszania;  $V$  – macierz reprezentująca błędy i zakłócenia;  $W$  – macierz separująca [4].

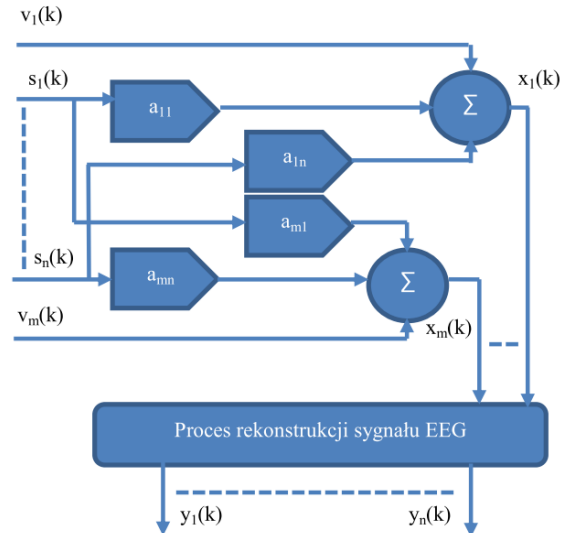
W równaniu (2) przedstawiono wektor sygnałów rejestrowanych, jako iloczyn wektorów nieznanych sygnałów źródłowych ze współczynnikami macierzy mieszającej w sumie z wektorem  $v(k)$  reprezentującym artefakty zakłócające:

$$x(k) = As(k) + v(k) \quad (2)$$

gdzie:  $k=1, 2, \dots, N$ ;  $x(k)=[x_1(k), \dots, x_m(k)]^T$  – wektor sygnałów rejestrowanych;  $s(k)=[s_1(k), \dots, s_n(k)]^T$  – wektor nieznanych sygnałów źródłowych.



Rys. 3. Schematyczne przedstawienie konstrukcji wektorów  $x(k)$  oraz  $y(k)$

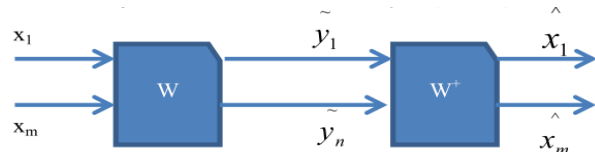


Rys. 4. Korelacja współczynników macierzy mieszającej z poszczególnymi wektorami sygnałów

## 3. Zastosowanie pseudoinwersji Moore-Penrose w rekonstrukcji sygnału EEG

W celu obiektywnej identyfikacji źródeł sygnałów, która jest niezmiernie ważna z punktu widzenia prowadzonych pomiarów, konieczne jest prawidłowe przedstawienie matematyczne ww. źródeł, jak i sygnałów odczytywanych na powierzchni głowy oraz użycie wysoce wyspecjalizowanych narzędzi w celu eliminacji tych składowych sygnałów, które są artefaktami. Zostało to potwierdzone w prowadzonych badaniach z których to wynika, że sygnał EEG można poddawać analizie przy pomocy pseudoinwersji Moore-Penrose [17].

Pseudoodwrotność została opracowana niezależnie przez dwóch naukowców Moore'a E. H. oraz Penrose'a R. [20]. Na rysunku 5 przedstawiono schematycznie rekonstrukcję rejestrowanych sygnałów źródłowych. Macierz separująca  $W$  jest równa pseudoinwersji Moore-Penrose z macierzy  $A$  (wzór 3).



Rys. 5. Rekonstrukcja sygnałów rejestrowanych

$$W = A^+ \quad (3)$$

gdzie:  $A^+$  – pseudoinwersja Moore-Penrose.

$$y(k) = Wx(k) \quad (4)$$

gdzie:  $y(k)$  – specyficzne komponenty sygnału.

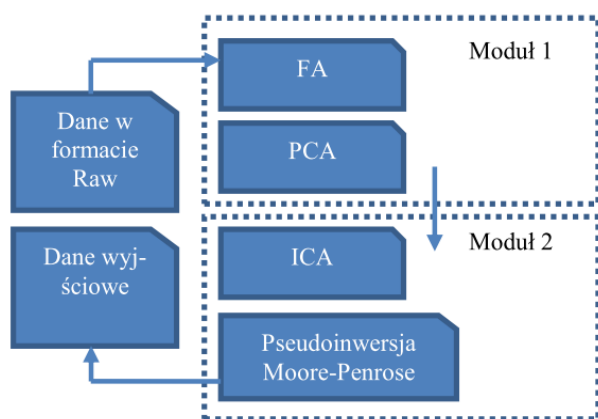
$$\hat{x}(k) = W^{-1} \tilde{y}(k) \quad (5)$$

gdzie:  $W^{-1}$  – macierz separująca po pseudoinwersji,

$\tilde{y}(k)$  – wektor powstały po usunięciu wszystkich artefaktów z  $y(k)$ . Z pseudoinwersją Moore-Penrose macierzy mamy do czynienia niezależnie od faktu, czy macierz jest kwadratowa, czy prostokątna.

#### 4. Dwumodułowy system analizy danych EEG

Podczas prowadzonych badań nad rozwojem interfejsów mózg-komputer zauważono potrzebę opracowania dwumodułowego systemu analizy danych EEG (rysunek 6), pracującego na bazie analizy czynnikowej – FA (ang. *Factor Analysis*) w tym metody PCA [14] oraz pseudoinwersji Moore-Penrose [13]. Niewątpliwie dzięki praktycznej implementacji proponowanego systemu jakość rekonstruowanego sygnału wyjściowego, który następnie może być z powodzeniem wykorzystany w procesach sterowania ulegnie polepszeniu [8]. PCA to metoda analizy składowych głównych zawarta w module 1 proponowanego systemu, natomiast metoda ICA (ang. *Independent Component Analysis*) to metoda analizy składowych niezależnych [12], w proponowanym systemie skorelowana z pseudoinwersją Moore-Penrose (moduł 2).



Rys. 6. System dwumodułowy łączący analizę czynnikową z pseudoinwersją Moore-Penrose

Przetwarzanie danych elektroencefalograficznych to proces wysoce złożony. Przede wszystkim w aspekcie uzyskania jak najwyższej jakości danych wyjściowych ze szczególnym uwzględnieniem źródeł ich powstawania. Dane w formacie wejściowym \*.raw, uzyskiwane są za pomocą dostępnych obecnie na rynku urządzeń do pomiarów elektroencefalograficznych, w tym z urządzeń komercyjnych oraz tych popularnych w zastosowaniach naukowych, firm takich jak: Emotiv Inc., NeuroSky etc.

#### 5. Wnioski

W artykule zaproponowano koncepcję realizacji przetwarzania sygnału biologicznego – EEG w oparciu o dwumodułowy system, w skład którego wchodzi analiza czynnikowa oraz metoda analizy składowych niezależnych, ze szczególnym uwzględnieniem pseudoinwersji Moore-Penrose. Obecnie wiele urządzeń pracujących na bazie technologii BCI (ang. *Brain Computer Interfaces*) dostarcza naukowcom niezliczoną ilość danych w postaci plików \*.raw. Dane te można z powodzeniem wykorzystywać do implementacji w realizacji aplikacji sterujących, zarówno: kursorem myszki na monitorze komputera, jak także pojazdami mobilnymi, czy też dronami [21-24]. Niezwykle ważna jest przy tym kwestia odpowiedniej analizy danych otrzymanych z elektrod urządzenia pomiarowego, a także poprawna identyfikacja źródeł mających swoje podstawy w mózgu osoby badanej.

#### Literatura

- Accardo A., Affinito M., Carrozzini M., Bouquet F.: Use of the fractal dimension for the analysis of electroencephalographic time series, *Biol. Cybern.*, vol. 77, 1997, 339-350.
- Bakardjian H., Cichocki A., Cincotti F., Mattia D., Babiloni F., Grazia Marciari M., De Vico Fallani F., Miwakeichi F., Yamaguchi Y., Martinez P., Salinari S., Tocci A., Astolfi L.: Estimate of causality between cortical spatial patterns during voluntary movements in normal subjects, *International Journal of Bioelectromagnetism* 8 (1), II/1-II/18, 2006.
- Bielińska E. et al.: *Identyfikacja Procesów*, Gliwice, Wydawnictwo Politechniki Śląskiej, 1997.
- Cheung Y.M., Xu L.: Dual multivariate auto-regressive modeling in state space for temporal signal separation, *IEEE T. Syst. Man. Cyb.* 33 2003, 386-398.
- Cichocki A., Zdunek R., Amari S.: *Csiszar's divergences for non-negative matrix factorization: Family of new algorithms*. LNCS 3889, Springer, 32-39.
- Cichocki A., Zdunek R., Amari S.: New algorithms for non-negative matrix factorization in applications to blind source separation. In *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP-2006*.
- Cruces S., Cichocki A., Castedo L.: An iterative inversion approach to blind source separation. *IEEE Trans. on Neural Networks*, 11 (6), 2000, 1423-1437.
- Cruces S.A., Castedo L., Cichocki A.: Robust blind source separation algorithms using cumulants, *Neurocomputing*, 49, 2002, 87-118.
- David O., Friston K.J.: A neural mass model for MEG/EEG: coupling and neuronal dynamics, *NeuroImage* 20 (3), 2003, 1743-1755.
- Dvorak I., Holden A.V.: *Mathematical Approaches to Brain Functioning Diagnostics*, Manchester Univ. Press, 1991.
- Gomez-Herrero G., De Clercq W., Anwar H., Egiastian K. Kara, Van Hu\_e O.S., Van Paesschen W.: Automatic removal of ocular artifacts in the eeg without a reference eeg channel, In *Proc. NORISIG*, Reykjavik, Iceland 2006, 130-133.
- Hyvarinen A., Kashunen J., Oja E.: *Independent Component Analysis*, John Wiley & Sons, Ltd, UK, 2001.
- Katsikis V.N., Pappas, D.: Fast computing of the Moore-Penrose inverse matrix, *Electronic Journal of Linear Algebra* 17(1), 2008, 637-650.
- Lagerlund T.D., Sharbrough F.W., Busacker N.E.: Spatial filtering of multichannel electroencephalographic recordings through principal component analysis by singular value decomposition, *J. Clin. Neurophysiol.*, vol. 14, 1997, 73-82.
- Lee D.D., Seung H. S.: Learning of the parts of objects by non-negative matrix factorization. *Nature*, 401, 1999, 788-791.
- Li Y., Cichocki A., Amari S.: Analysis of sparse representation and blind source separation. *Neural Computation*, 16 (6), 2004, 1193-1204.
- Li Y., Cichocki A., Amari S.: Blind estimation of channel parameters and source components for EEG signals: A sparse factorization approach. *IEEE Transactions on Neural Networks*, 2006, 17, 419-431.
- Li Y., Cichocki A., Amari S., Shishkin S., Cao J., Gu F.: Sparse representation and its applications in blind source separation. In *Seventeenth Annual Conference on Neural Information Processing Systems (NIPS-2003)*. Vancouver.
- Lin C.J.: *Projected gradient methods for non-negative matrix factorization* (Tech. Rep.) Department of Computer Science, National Taiwan University, 2005.
- Petralias A., Katsikis V.N., Pappas D.: An improved method for the computation of the Moore-Penrose inverse matrix, *Applied Mathematics and Computation* 217(23) 2011, 9828-9834.
- Paszkiel S.: Augmented reality of technological environment in correlation with brain computer interfaces for control processes, *Advances in Intelligent Systems and Computing* 267 - AISC, Springer Switzerland 2014, 197-203.
- Paszkiel S.: The use of Brain Computer Interfaces in the control processes based on industrial PC in terms of the methods of EEG signal analysis, *Journal of Medical Informatics & Technologies - Vol. 22* 2013, 55-62.
- Paszkiel S., Blachowicz A.: The application of electroencephalographic signals in the aspect of controlling a mobile robot for measurements of incomplete discharges, *Przegląd Elektrotechniczny*, R. 86 NR 8/2010, 303-306.
- Paszkiel S.: The population modeling of neuronal cell fractions for the use of controlling a mobile robot. *Pomiary, Automatyka, Robotyka*, vol. 2, 2013, 254-259.

Dr inż. Szczepan Paszkiel  
e-mail: s.paszkiel@po.opole.pl



Doktor inżynier Szczepan Paszkiel, pracuje obecnie na stanowisku adiunkta w Instytucie Automatyki i Informatyki Politechniki Opolskiej. Tytuł doktora nauk technicznych obronił w 2011 r. z wyróżnieniem. Zainteresowanie naukowe doktora skupiają się wokół automatyki, informatyki oraz pomiarów biomedycznych.

Autor ponad 80 publikacji naukowych oraz jednego z pierwszych w Polsce podręczników na temat neuroinformatyki. Uczestnik wielu konferencji naukowych w kraju i zagranicą. Członek Komisji Metrologii Polskiej Akademii Nauk. Ekspert naukowy Narodowego Centrum Badań i Rozwoju.

otrzymano/received: 2014.09.19

przyjęto do druku/accepted: 2014.10.14