



Klasyfikacja słuchowych potencjałów wywołanych w oparciu o dekompozycję falkową i sieć SVM

MICHAŁ SUCHOCKI, ANDRZEJ DOBROWOLSKI,
EWELINA MAJDA-ZDANCEWICZ, KAZIMIERZ TOMCZYKIEWICZ¹

Wojskowa Akademia Techniczna, Wydział Elektroniki, 00-908 Warszawa, ul. gen. S. Kaliskiego 2,
Michal.Suchocki@gmail.com, Andrzej.Dobrowolski@wat.edu.pl, Ewelina.Majda@wat.edu.pl

¹Wojskowy Instytut Medyczny, Klinika Neurologiczna, 04-141 Warszawa, ul. Szaserów 128,
KTomczykiewicz@wim.mil.pl

Streszczenie. W elektrofizjologicznej ocenie słuchu oraz diagnozowaniu uszkodzeń pnia mózgu najczęściej wykorzystuje się słuchowe potencjały wywołane pnia mózgu o krótkiej latencji. Charakteryzują się one kolejno ułożonymi w funkcji czasu maksimami, zwanymi *załamkami* lub *falami*. Morfologia przebiegu, a w szczególności zależności czasowe i amplitudowe poszczególnych załamków, umożliwiają neurologowi postawienie diagnozy, co nie jest zadaniem prostym. Wymaga od neurologa doświadczenia, skupienia uwagi i bardzo dobrej percepcji. W celu wsparcia procesu diagnostycznego autorzy opracowali algorytm realizujący zautomatyzowaną klasyfikację słuchowych potencjałów wywołanych do grupy przypadków patologicznych i fizjologicznych, z czułością i specyficznością określoną na niezależnej grupie testowej liczącej 50 przypadków, wynoszącą odpowiednio 84% i 88%.

Słowa kluczowe: inżynieria biomedyczna, słuchowe potencjały wywołane pnia mózgu, dekompozycja falkowa, sieć wektorów nośnych

DOI: 10.5604/12345865.1186236

Wstęp

Przewodzenie sygnałów poprzez włókna nerwowe odbywa się dzięki złożonym procesom elektrochemicznym, a każdemu sygnałowi przenoszonemu przez nerw odpowiada tzw. potencjał czynnościowy, który można indywidualnie rejestrować, stosując właściwe techniki elektroneurograficzne. Do celów diagnostycznych wykonuje się badania elektroencefalograficzne — EEG, które pozwalają na ocenę

funkcjonowania dużych zespołów włókien nerwowych oraz wybranych ośrodków korowych mózgu [1]. Szczególnym rodzajem badań EEG są badania tzw. potencjałów wywołanych — PW (ang. *EP* — *Evoked Potentials*). Badania te polegają na rejestracji napięcia pomiędzy elektrodami umieszczonymi w różnych punktach na powierzchni głowy, po zadziałaniu fizycznych bodźców drażniących receptory słuchowe, wzrokowe, czuciowe lub — rzadziej — węchowe. Badania obejmują rejestrację odpowiedzi od receptorów poprzez poszczególne fragmenty traktu nerwowego, aż po ośrodki korowe mózgu. W ramach niniejszego artykułu autorzy skoncentrowali się na słuchowych potencjałach wywołanych — SPW (ang. *AEP* — *Auditory Evoked Potentials* lub ABR — *Auditory Brainstem Response*) [2].

Ucho, m.in. ze względu na swą wyjątkowo skomplikowaną budowę mechaniczną, uznawane jest za najbardziej złożony zmysł człowieka. Tor przenoszenia fali akustycznej rozpoczyna się w małżowinie usznej. Następnie, poprzez kanał słuchowy i błonę bębenkową, prowadzi do układu kosteczek: młoteczka, kowadełka i strzemiączka, dopasowujących impedancję powietrznego środowiska transmisyjnego do płynnego środowiska ślimaka tak, aby fala dźwiękowa rozchodząca się w powietrzu mogła efektywnie pobudzić cieczę ślimaka. Zakończeniem toru przenoszenia jest trąbka Eustachiusza i okienko owalne. Część sensoryczną stanowi narząd spiralny, z którego zakodowana informacja przenoszona jest drogą słuchową obejmującą nerw słuchowy, pień mózgu, ciało kolankowate przyśrodkowe, promienistość słuchową i ostatecznie płaty skroniowe kory mózgu, gdzie dźwięk jest właściwie interpretowany.

Nawet przy silnych bodźcach akustycznych potencjały wywołane rejestrowane na powierzchni głowy mają bardzo niskie poziomy w stosunku do innych sygnałów o charakterze szumowym, dlatego w celu ich wyizolowania stosuje się technikę synchronicznego uśredniania bardzo dużej liczby odpowiedzi na sekwencję bodźców akustycznych. Badania elektrofizjologiczne słuchu polegają więc na synchronicznej rejestracji napięć powstałych wskutek zadziałania bodźca słuchowego poprzez elektrody odpowiednio umieszczone na powierzchni głowy. W zależności od czasu opóźnienia odpowiedzi względem wystąpienia bodźca rozróżnia się potencjały z części obwodowej narządu słuchu, z pnia mózgu oraz potencjały korowe. Najszybciej, w czasie od ok. zera do 2 ms, pojawiają się sygnały mikrofoniczne ślimaka, potencjały sumacyjne oraz potencjał czynnościowy nerwu słuchowego. Diagnostykę tych potencjałów nazywa się elektrochleografią (*ECoG*). W dalszej kolejności występują potencjały z pnia mózgu — ABR (ang. *Audiometry Brainstem Responses*), których opóźnienie (latencja) wynosi od ok. 2 ms do 9 ms (zwane potencjałami wczesnymi). Dalsze piętra układu nerwowego reprezentowane są przez specyficzne dla bodźców akustycznych potencjały o latencji od 12 ms do 50 ms oraz potencjały korowe (90-200 ms) łącznie z kognitywnym potencjałem P300 [1]. Z punktu widzenia przydatności klinicznej, w badaniu słuchu, diagnozowaniu uszkodzeń pnia mózgu oraz innych schorzeń neurologiczno-otologicznych najistotniejsze są słuchowe

potencjały wywołane pnia mózgu — SPWPM (BAEP — ang. *Brainstem Auditory Evoked Potentials*) [3]. W otologii badanie SPWPM wykorzystuje się do określenia progu słyszalności i różnicowania ubytków słuchu, natomiast w neurologii umożliwia ono rozpoznanie i lokalizację pozaślimakowych uszkodzeń drogi słuchowej. Dzięki SPWPM można potwierdzić guzy kąta mostowo-mózdkowego, śmierć pnia mózgu, stwardnienie rozsiane, a także monitorować operacje neurochirurgiczne w celu wykluczenia zagrożenia uszkodzenia drogi słuchowej.

Przebiegi czasowe SPWPM charakteryzują się kolejno ułożonymi w funkcji czasu maksimami, zwanymi w środowisku lekarskim załamkami, przy czym ze względu na stałość i powtarzalność ocenia się głównie załamki I, III i V [3]. Występowanie kolejnych załameków wynika z anatomicznej budowy drogi słuchowej człowieka. Latencje poszczególnych załameków, ich amplitudy, stosunki amplitud, interlatencje oraz generalnie występowanie bądź nieobecność w określonym segmencie przebiegu czasowego umożliwiają neurologowi odpowiednią klasyfikację. Okazuje się jednak, że ze względu na różnorodność rejestrowanych przebiegów nie jest to zadanie łatwe i wymaga od neurologa dużego doświadczenia, skupienia uwagi i bardzo dobrej percepcji.

W celu wsparcia procesu diagnostycznego autorzy opracowali algorytm realizujący zautomatyzowaną klasyfikację słuchowych potencjałów wywołanych do grupy potencjałów fizjologicznych lub patologicznych. Opracowanie finalnego oprogramowania diagnostycznego objęło procedury ekstrakcji przebiegów czasowych SPWPM z plików binarnych, generacji i selekcji cech dystynktywnych oraz implementację skutecznego klasyfikatora.

1. Repozytorium słuchowych potencjałów wywołanych pnia mózgu

W pierwszym etapie badań zbudowano bazę danych zawierającą przebiegi słuchowych potencjałów wywołanych pnia mózgu zarejestrowane w grupie osób zdrowych i chorych z objawami uszkodzenia słuchu lub pnia mózgu o różnej etiologii. Następnie dokonano wstępnej analizy w celu opracowania optymalnej metody generacji cech dystynktywnych. Poszukiwane cechy powinny umożliwiać jak najwierniejszą rekonstrukcję sygnału z jak najmniejszego ich zbioru, a jednocześnie charakteryzować się jak największą zdolnością dyskryminacyjną.

W celu pozyskania wystarczającej liczby rejestracji słuchowych potencjałów wywołanych przeprowadzono szereg badań klinicznych zgodnych z obowiązującymi normami medycznymi. Badanie odbywało się w pozycji siedzącej, a badanego proszono o zrelaksowanie się. Osobie badanej zakładano słuchawki stereofoniczne oraz w odpowiednich miejscach, tj. na wyrostkach sutkowatych i na czubku głowy, mocowano elektrody powierzchniowe. Bodziec akustyczny miał charakter impulsowy,

a jego brzmienie można utożsamiać z serią kliknięć przycisku myszy komputerowej. Pojedynczy bodziec wytwarzany był przez prostokątny impuls elektryczny o czasie trwania 100 μ s. Z uwagi na pomiar wczesnych odpowiedzi wywołanych, bodziec był powtarzany z częstotliwością 11,4 Hz. Bodźcem dla drugiego ucha był szum maskujący o stałym natężeniu. Poziom głośności stymulacji wahał się w granicach 70 dB powyżej progu słyszalności przy 30 dB poziomie szumu maskującego. Badanie drugiego ucha wykonywano, przełączając akustyczny bodziec stymulujący i szum, bez zmiany położenia elektrod.

Ponieważ odpowiedzi na pojedyncze bodźce charakteryzują się bardzo małą amplitudą, dlatego bodźce powtarza się wielokrotnie, a odpowiedź uśrednia synchronicznie. Liczba powtórzeń bodźca w serii wynosiła 2000, a więc pojedyncza rejestracja trwała około 3 min.

Zgodnie z zaleceniami IFCN (ang. *International Federation of Clinical Neurophysiology*) do badania zastosowano trzy elektrody: dwie odbiorcze $A_{stym.}$ (po stronie stymulacji) i $A_{szum.}$ (po stronie maskowania szumem) oraz jedną odniesienia $A_{odnie.}$. Elektrody odbiorcze umocowane były na wyrostkach sutkowych, bądź — rzadziej — na małżowinach usznych. Elektrode odniesienia umieszczono na szczycie głowy. Z zamocowanych na głowie pacjenta elektrod utworzono w ten sposób następujące pary: stymulującą $A_{stym.}$ — $A_{odnie.}$ i maskującą $A_{szum.}$ — $A_{odnie.}$. Para stymulująca używana była do rejestrowania odpowiedzi na generowane „kliknięcia”, natomiast para maskująca rejestrowała odpowiedź układu słuchowego na szum [4]. W ramach automatycznych procedur opracowywanych przez autorów analizowano sygnał rejestrowany przez parę stymulującą. W przypadku wątpliwości diagnostycznych neurolog dokonywał weryfikacji, polegającej na uwzględnieniu wyników rejestracji dokonanej również przez parę maskującą.

Konfiguracja urządzenia pomiarowego oraz elektrod umożliwiała neurologowi poprawne określenie załamek przy czasie analizy równym 10 ms (zastosowano cyfrowy dwubiegunowy filtr górnoprzepustowy z dolną częstotliwością graniczną wynoszącą 150 Hz oraz dolnoprzepustowy filtr jednobiegunowy z górną częstotliwością graniczną wynoszącą 3 kHz). Wartości graniczne norm różnic między latencjami poszczególnych załamek w zaleceniach IFCN, w populacji zdrowych dorosłych ludzi, przedstawiono w tabeli 1 [3], a zobrazowanie widziane przez lekarza na monitorze systemu na rysunku 1.

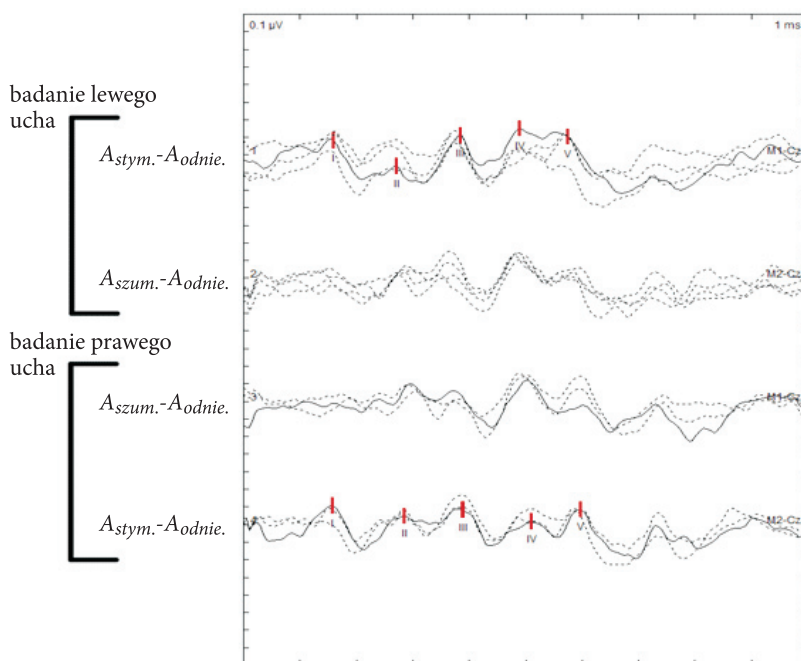
TABELA 1

Latencje załamek według norm IFCN [3]

Różnica latencji między załamekami	I-III	III-V	I-V
Górny limit [ms]	2,5	2,4	4,5

W ocenie SPWPM zasadniczo uwzględnia się: obecność i latencje załameków I, III i V, interlatencję między załamekami I-III, III-V i I-V oraz stosunek wartości maksymalnych załameków V i I [3]. W procesie diagnostycznym porównuje się również latencje załameków określonych dla lewego i prawego ucha, a w celu ustalenia progu słyszalności dokonuje się regulacji bodźca akustycznego, tak by przy jego minimalnej wartości otrzymać załamek V. Każde laboratorium powinno dokonać kalibracji progów słyszalności wg norm określonych przez IFCN.

Obecnie diagnostyka w oparciu o przebiegi SPWPM polega na wzrokowej interpretacji zarejestrowanego przez lekarza przebiegu, z kontrolą liczbowych wartości latencji i interlatencji oraz stosunków amplitud. W praktyce diagnosta niejednokrotnie musi powtarzać badanie w celu uzyskania odpowiedniego, nadającego się do oceny przebiegu. Stwierdzenie patologii bądź poprawności jest więc dość subiektywne, a istotną rolę odgrywa doświadczenie i zdolności percepcyjne osoby wykonującej badanie.



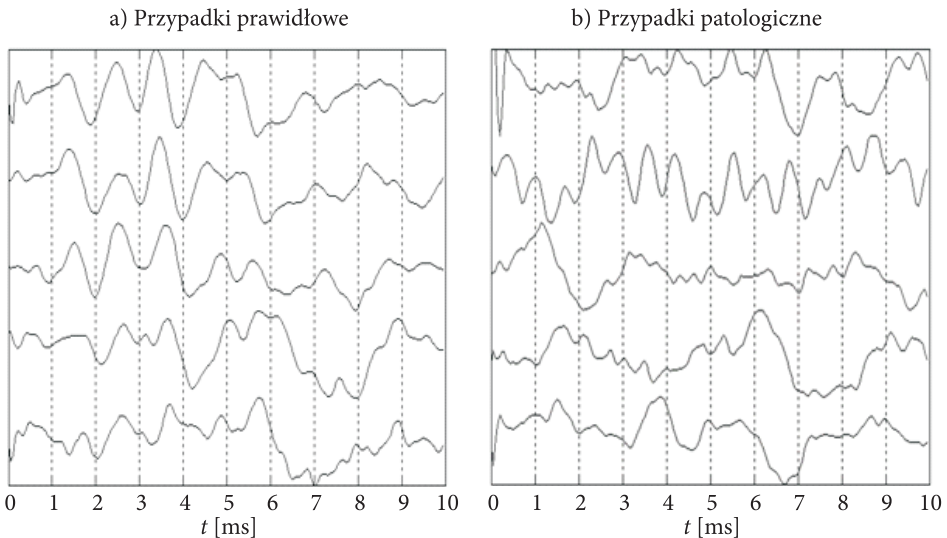
Rys. 1. Słuchowe potencjały wywołane pnia mózgu obserwowane przez lekarza podczas badania; markerami oznaczono kolejne załameki

Autorzy artykułu podjęli się zadania opracowania automatycznego analizatora przebiegów SPWPM, bazującego na surowych danych otrzymywanych bezpośrednio z karty pomiarowej, bez naniesionych przez diagnostę markerów oznaczających

poszczególne załamki. Poważną przeszkodą na tym etapie okazały się urządzenia medyczne rejestrujące potencjały wywołane i zapisujące je do plików binarnych o formacie znanym wyłącznie producentowi. Bezpośredni wgląd w bazę danych pacjentów możliwy jest jedynie poprzez graficzny interfejs użytkownika-diagnosty. Producenci nie udostępniają dodatkowych narzędzi umożliwiających zapis surowych danych w postaci np. plików tekstowych w formacie ASCII. Ze względu na prawa autorskie i ochronę własnych interesów firmy bardzo niechętnie udostępniają kod źródłowy programu, a udostępniana część struktury danych może być niekompletna i mało użyteczna w procesie wydobywania danych i konwersji do formatu otwartego. Mimo poważnych problemów z pozyskaniem informacji o sposobie zapisu danych, ostatecznie została jednak opracowana wtyczka programowa (*plug-in*) umożliwiająca konwersję danych z plików binarnych systemu *Nicolet Viking Select*.

Badania przeprowadzono na grupie 130 osób obu płci, w wieku od 19 do 86 lat. W oparciu o diagnozę postawioną przez neurologa grupę podzielono na dwie podgrupy: prawidłową i patologiczną. Z każdego badania EEG pozyskano maksymalnie szesnaście przebiegów AEP. W badaniach wykorzystano dane wyłącznie tych osób, które wyraziły chęć i zgodę na ich przeprowadzanie. Wszelkie dane osobowe uzyskane w trakcie badania są poufne i nie będą w żadnym wypadku ani w żadnej formie ujawniane innym osobom ani podmiotom.

Wśród wszystkich przebiegów zarejestrowanych w opisanym wcześniej procesie diagnostycznym znalazło się po 65 sygnałów reprezentujących przypadki prawidłowe i patologiczne (po jednym sygnale pochodzącym od każdej osoby). Z tak otrzymanego zbioru wylosowano próbę uczącą (po 40 sygnałów z każdej grupy) oraz



Rys. 2. Przykładowe przebiegi słuchowych potencjałów wywołanych: a) prawidłowe i b) patologiczne

próbę testującą (po 25 sygnałów z każdej grupy). W celu ilustracji różnic między przebiegami na rysunku 2 przedstawiono po pięć reprezentatywnych przebiegów z grupy prawidłowej i patologicznej.

2. Generacja i selekcja cech dystynktywnych

Opierając się na wcześniejszych badaniach [5-8], ekstrakcji cech dokonano za pomocą dyskretnej dekompozycji falkowej (ang. DWT — *Discrete Wavelet Transform*) [9], natomiast klasyfikację przebiegów przeprowadzono przy użyciu liniowej sieci SVM (ang. *Support Vector Machine*) [10]. Efektywność takiego podejścia, w kontekście selekcji pojedynczych przebiegów, została potwierdzona w [11, 12]. Ze względu na szybkość wyznaczania aproksymacji i detali oraz bezproblemową rekonstrukcję sygnału, umożliwiającą bieżącą kontrolę działania algorytmu, zdecydowano się na dekompozycję falkową realizowaną szybkim algorytmem Mallata [13] na bazie falek ortogonalnych o zwartym nośniku. Do wstępnej selekcji wybrano falki *Daubechies*, *Symlet* i *Coiflet*. W iteracyjnym procesie optymalizacyjnym — korzystając z wielu metod eksploracji danych, m.in. z kryterium Fishera, obrazowania PCA (ang. *Principal Component Analysis*) oraz analizy LDA (ang. *Linear Discriminant Analysis*) — opracowano sposób synchronizacji przebiegów, zoptymalizowano typ i rząd użytej falki oraz liczbę poziomów dekompozycji, a także dokonano selekcji cech.

Podlegający ocenie diagnostycznej przebieg słuchowego potencjału wywołanego pnia mózgu, o czasie trwania 10 ms, liczy 200 próbek pobieranych z szybkością 20 tys. próbek na sekundę. Kluczowym rozwiązaniem, które znacząco poprawiło osiągnięte rezultaty, było zastosowanie analizy w przedziałach czasowych skoncentrowanych wokół pierwszych pięciu załamek. Analizując sygnały wyjściowe kilku różnych testowych klasyfikatorów (PCA, LDA i SVM) oceniono, że wystarczającą informację dyskryminacyjną niosą detale na czterech pierwszych poziomach dekompozycji przy zastosowaniu falki *Daubechies* piątego rzędu, co odpowiada pasmu sygnału 625-10000 Hz.

Po przeprowadzeniu szeregu eksperymentów okazało się, że z przedziałów skoncentrowanych wokół jedynie czterech pierwszych załamek wystarczy wyekstrahować po dwa współczynniki falkowe o największych wartościach z każdego poziomu, co pozwoliło utworzyć zbiór 32 deskryptorów numerycznych opisujących każdy z przebiegów SPWPM. Zdefiniowane na etapie generacji cech deskryptory stanowią maksymalny zbiór potencjalnych cech dystynktywnych, które mogą być wykorzystane w procedurze klasyfikacji. Badania prowadzone na świecie pokazują, że nie zawsze użycie maksymalnego zestawu cech prowadzi do najlepszych wyników, gdyż nie są one jednakowo ważne w procesie rozpoznania wzorców. Wśród wyekstrahowanych cech pewne mogą mieć postać szumu pomiarowego pogarszającego możliwości poprawnej klasyfikacji, natomiast inne cechy są silnie skorelowane i mogą zwykle

niekorzystnie wpływać na jakość klasyfikacji, dominując nad innymi i tłumiąc w ten sposób ich korzystny wpływ [14]. Nie oznacza to jednak, że zmienne te są źle określone, pozbawione jakiejkolwiek wartości czy błędnie zmierzone; są one po prostu niezwiązane z rozpatrywanym w danym momencie zadaniem. W takim przypadku mówi się o braku korelacji z klasą, a cechy takie są często nazywane szumowymi. Ważnym elementem procesu staje się zatem właściwa selekcja cech, pozwalająca na utworzenie optymalnego wektora wyekstrahowanych deskryptorów.

W badaniu jakości cech można wyróżnić dwie strategie. W pierwszej bada się każdą cechę niezależnie od zastosowanej metody klasyfikacji, oceniając jej jakość pod kątem różnicowania klas bez uwzględnienia konkretnego klasyfikatora. Druga to selekcja oparta na ścisłej współpracy z klasyfikatorem. Autorzy zastosowali równoległe obie strategie, starając się w końcowej fazie selekcji zapewnić minimalizację liczby wektorów nośnych w klasyfikatorze SVM bez pogarszania wyników klasyfikacji w zbiorze uczącym. Jedną z głównych zalet stosowania sieci SVM jest uzyskanie klasyfikatora o największej zdolności generalizacji — przyjmuje się, że sieć SVM w roli klasyfikatora w zdecydowanej większości zadań jest obecnie bezkonkurencyjna [14].

3. Klasyfikacja

W ostatnim etapie badań dokonano selekcji cech i ustalono ich ostateczny zbiór, który posłużył do budowy finalnego klasyfikatora. W trakcie badań monitorowano miarę Vapnika-Chervonenkisa $VCdim$, która została zdefiniowana jako liczebność n największego zbioru danych wzorców, dla których system może zrealizować wszystkie możliwe 2^n dychotomii tego zbioru (podziału zbioru na dwie części za pomocą prostej). $VCdim$ odgrywa istotną rolę przy określaniu minimalnej liczby danych uczących p , gdyż dla uzyskania dobrej generalizacji musi być spełniony warunek

$$p \gg VCdim. \quad (1)$$

Według Vapnika podczas klasycznego uczenia sieci neuronowej dobre rezultaty generalizacji otrzymuje się zawsze, jeżeli liczba danych uczących jest co najmniej 20 razy większa niż miara $VCdim$. Oczywiście bywa, że udaje się to przy liczbie p dużo mniejszej. W przypadku sieci SVM:

$$VCdim \leq \min \left(\left\lceil \frac{D^2}{\rho^2} \right\rceil, N \right) + 1, \quad (2)$$

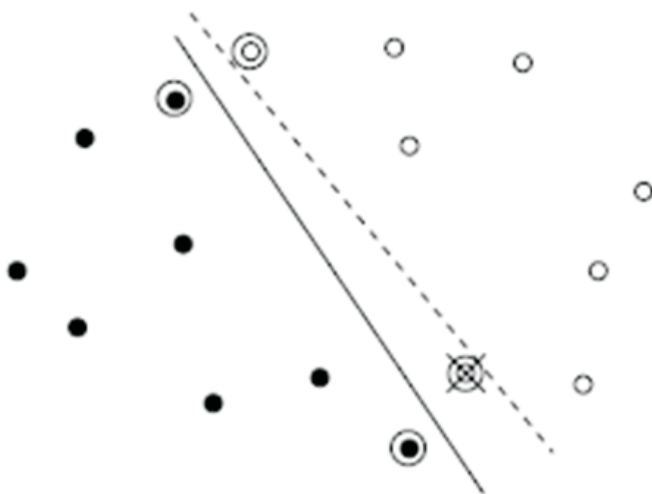
gdzie: N — liczba neuronów ukrytych (wymiar wektora cech);

$\lceil \cdot \rceil$ — funkcja *entier* — zwraca najmniejszą liczbę całkowitą większą/równą od argumentu;

D — średnica najmniejszej kuli w przestrzeni N -wymiarowej obejmującej wszystkie wektory uczące (tj. p);
 ρ — szerokość marginesu separacji.

Powyższa nierówność potwierdza, że możliwe jest sterowanie wartością miary $VCdim$, niezależniąc ją od wymiaru N przestrzeni cech poprzez zastosowanie szerokiego marginesu [14].

Dla sieci SVM w roli klasyfikatora szacuje się często górną granicę prawdopodobieństwa błędnej klasyfikacji jako stosunek liczby wektorów nośnych do ogólnej liczby danych uczących pomniejszonej o jeden; stąd dla zwiększenia zdolności generalizacyjnych dąży się do zmniejszenia liczby wektorów nośnych nawet kosztem zwiększenia liczby błędnych klasyfikacji w zbiorze uczącym [14]. W celu interpretacji tej zależności rozpatrzmy typową sytuację po treningu z wykorzystaniem zbioru uczącego o liczebności p , przedstawioną na rysunku 3. Wektory nośne (kółka) mogą stać się błędami (skreślone kółko) po wykluczeniu i ponownym wytrenowaniu — linia przerywana oznacza nową granicę decyzyjną [10].



Rys. 3. Ilustracja wpływu wektorów nośnych [10]

Korzystając z metody krosvalidacji w wariancie *Leave-One-Out*, można oszacować błąd testowania poprzez usunięcie jednego z punktów uczących, ponowne wytrenowanie sieci na pozostałych $(p - 1)$ punktach oraz przetestowanie na punkcie usuniętym, a następnie iteracyjne powtórzenie uczenia i testowania dla wszystkich punktów uczących. Z zasady działania sieci SVM wynika, że usuwanie dowolnych punktów uczących, które nie są wektorami nośnymi, nie ma wpływu na położenie hiperpłaszczyzny decyzyjnej oraz że w najgorszym razie może się zdarzyć, że każdy wektor nośny będzie związany z błędnie sklasyfikowanym punktem.

Liczba błędnych klasyfikacji wektora testowego w pojedynczej — i -tej — iteracji może być równa zero, gdy wykluczony i testowany jest wektor niebędący wektorem nośnym wyznaczonym na pełnym zbiorze uczącym, albo zero lub jeden, w przypadku gdy wykluczony został wektor będący wektorem nośnym wyznaczonym na pełnym zbiorze uczącym, a zatem sumaryczna liczba błędów we wszystkich p iteracjach nie przekroczy liczby wektorów nośnych wyznaczonych na pełnym zbiorze uczącym. Odniesienie takiej liczby błędów do pełnej liczby przykładów uczących dałoby jednak obciążony estymator błędu testowania. Ponieważ w istocie sieć trenowana jest zbiorem o liczebności $(p - 1)$, lepszym oszacowaniem jest popularna i prosta zależność [15]

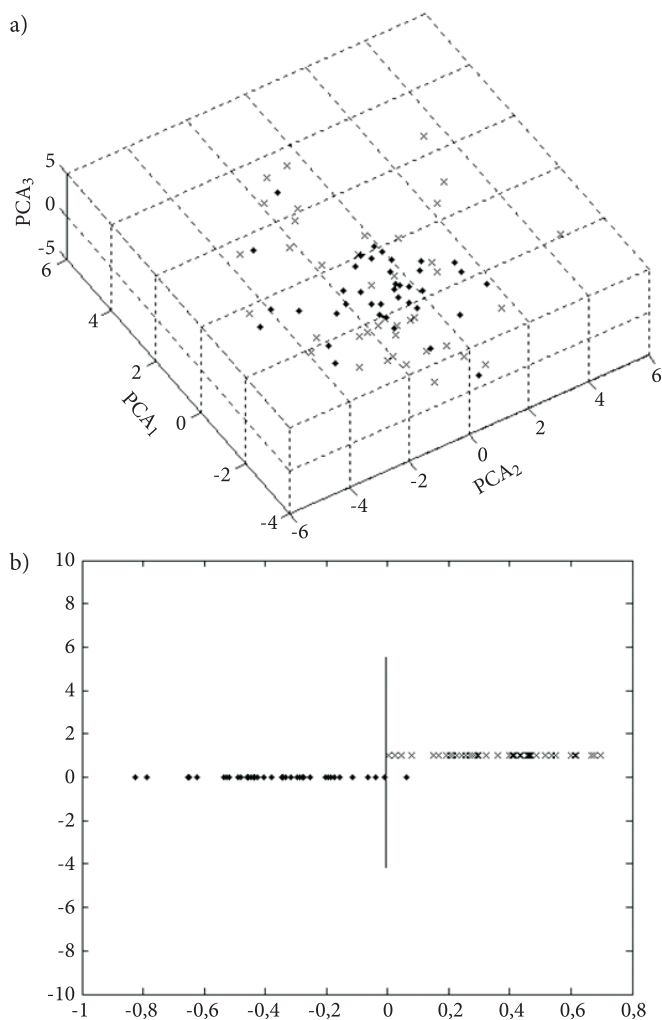
$$E_{\max} = \frac{N_{svp}}{p - 1}, \quad (3)$$

gdzie N_{svp} oznacza liczbę wektorów nośnych określoną na pełnym p -elementowym zbiorze uczącym.

Zasadniczym wnioskiem płynącym z zależności (3) jest stwierdzenie, że zmniejszanie liczby wektorów nośnych zmniejsza górne oszacowanie błędu klasyfikacji na danych testujących. W związku z tym, w celu poprawy generalizacji sieci SVM, należy minimalizować liczbę wektorów nośnych, nawet kosztem ograniczenia liczby prawidłowych klasyfikacji na zbiorze uczącym. Traktując ten fakt jako kryterium optymalizacji, autorzy zredukowali liczbę cech do 12, a liczbę wektorów nośnych do 16, utrzymując zerowy błąd klasyfikacji w grupie uczącej i potwierdzając jednocześnie liniową separowalność klas. Górne oszacowanie błędu klasyfikacji otrzymanego klasyfikatora SVM na danych testujących zgodnie z (3) wynosi więc 20%.

Korzystając z opisanego w rozdziale 1 zbioru testowego, autorzy przetestowali klasyfikator na niezależnej grupie testowej i w wyniku klasyfikacji zarejestrowali trzech niewłaściwie zaklasyfikowanych pacjentów zdrowych oraz czterech pacjentów chorych z nierozpoznanym schorzeniem, otrzymali więc błąd całkowity równy 14% oraz czułość 84% i specyficzność 88%.

W celu ilustracji otrzymanego rozwiązania, na rysunku 4 przedstawiono zobrażowanie wyekstrahowanych 12-wymiarowych wektorów reprezentujących przebiegi słuchowych potencjałów wywołanych z grupy uczącej w przestrzeni trzech najważniejszych składników głównych PCA oraz po przekształceniu LDA.



Rys. 4. Rzuty wyekstrahowanych wektorów cech w przestrzeni trzech pierwszych składowych PCA (a) oraz w zobrazowaniu LDA (b). Wypełnionymi kwadratami oznaczono przypadki prawidłowe, a znakami „x” patologiczne

Podsumowanie

W celu wsparcia procesu diagnostycznego opracowano algorytm realizujący zautomatyzowaną klasyfikację słuchowych potencjałów wywołanych. Do generacji zbioru cech dystynktywnych zastosowano falkową dekompozycję słuchowych potencjałów wywołanych pnia mózgu w odpowiednio dobranych przedziałach czasowych. Korzystając z wielu metod eksploracji danych (kryterium Fishera, PCA i LDA oraz sieć SVM), skonstruowano klasyfikator oparty na liniowej sieci SVM, charakteryzujący

się czułością na poziomie 84% i swoistością (specyficznością) na poziomie 88%, określonymi na niezależnej grupie testowej liczącej 50 przypadków.

W pełnym 130-osobowym zbiorze przypadków dokładność opracowanej metody wynosi 95%, a czułość i specyficzność odpowiednio 94% i 95%. Dla porównania, w [16] autorzy w oparciu o zbiór 85-osobowy (cechy zdefiniowano w dziedzinie czasu oraz w oparciu o dekompozycję falkową) przetestowali kilka różnych klasyfikatorów i największą dokładność równą 83,4% oraz specyficzność równą 86,3% uzyskali dla naiwnego klasyfikatora bayesowskiego. Z kolei największą czułość równą 93,1% zapewniła klasyfikacja za pomocą sieci SVM. Autorzy zaprezentowanej w artykule metody uzyskali więc bardzo dobre wskaźniki jakości.

Otrzymany w wyniku przeprowadzonych badań algorytm, po odpowiedniej adaptacji i rozbudowie, może znaleźć zastosowanie również przy analizie wzrokowych i somatosensorycznych potencjałów wywołanych, co stanowi główny wątek dalszych badań planowanych przez autorów artykułu. Autorzy planują również rozszerzenie funkcjonalności oprogramowania o predykcję konkretnych jednostek chorobowych powodujących zaburzenia morfologii przebiegu słuchowych potencjałów wywołanych pnia mózgu.

Wyniki przedstawione w artykule są efektem badań finansowanych w ramach działalności statutowej Wydziału Elektroniki Wojskowej Akademii Technicznej.

Artykuł opracowany na podstawie referatu wygłoszonego na IX Krajowej Konferencji „Diagnostyka Techniczna Urządzeń i Systemów” (DIAG’15), Ustroń, 22-25.09.2015 r.

Artykuł wpłynął do redakcji 10.07.2015 r. Zweryfikowaną wersję po recenzjach otrzymano 18.10.2015 r.

LITERATURA

- [1] JAROSZYK F. (red.), *Biofizyka*, Wydawnictwo Lekarskie PZWL, Warszawa, 2011.
- [2] VALDERRAMA J.T., DE LA TORRE A., ALVAREZ I., SEGURA J.C., THORNTON A.R.D., SAINZ M., VARGAS J.L., *Automatic quality assessment and peak identification of auditory brainstem responses with fitted parametric peaks*, Computer Methods and Programs in Biomedicine, vol. 114, 2014, pp. 262-275.
- [3] CHIAPPA K.H., *Evoked Potentials in Clinical Medicine*, Reven Press, New York, 1990.
- [4] BINNIE C.D., COOPER R., MAUGUIERE F., OSSELTON J., PRIOR P.F., TEDMAN M., *Clinical Neurophysiology: EMG, Nerve Conduction and Evoked Potentials*, vol. 1, 2nd Edition, Elsevier B.V, 2004.
- [5] DOBROWOLSKI A., SUCHOCKI M., TOMCZYKIEWICZ K., *Computer analysis of auditory brainstem evoked potentials*, Przegląd Elektrotechniczny, 87, nr 9a, 2011, pp. 145-150.
- [6] DOBROWOLSKI A., SUCHOCKI M., TOMCZYKIEWICZ K., *Digital signal processing in the diagnosis of brainstem auditory evoked potentials*, Elektronika: konstrukcje, technologie, zastosowania, vol. 53, nr 5, 2012, pp. 26-29.
- [7] SUCHOCKI M., DOBROWOLSKI A., *Obiektywna ocena traktu słuchowego oparta na analizie falkowej potencjałów wywołanych i sieci wektorów nośnych*, Przegląd Elektrotechniczny, 89, nr 9, 2013, s. 160-164.

- [8] DOBROWOLSKI A., WIERZBOWSKI M., TOMCZYKIEWICZ K., *Multiresolution MUAPs decomposition and SVM-based analysis in the classification of neuromuscular disorders*, Computer Methods and Programs in Biomedicine, vol. 107, no. 3, 2012, pp. 393-403.
- [9] DAUBECHIES I., *The wavelet transform, time-frequency localizations and signal analysis*, IEEE Trans. on Information Theory, vol. 36, no. 5, 1990, pp. 961-1005.
- [10] BURGESS C.J.C., *A tutorial on support vector machines for pattern recognition*, Data mining and knowledge discovery, vol. 2, 1998, pp. 121-167.
- [11] ACIR N., ÖZDAMAR Ö., GÜZELIS C., *Automatic classification of auditory brainstem responses using SVM-based feature selection algorithm for threshold detection*, Engineering Applications of Artificial Intelligence, vol. 19, 2006, pp. 209-218.
- [12] ACIR N., ERKAN Y., BAHTIYAR Y.A., *Auditory brainstem response classification for threshold detection using estimated evoked potential data: comparison with ensemble of averaged data*, Neural Computing and Applications, vol. 22, 2013, 859-867.
- [13] MALLAT S.G., *A Wavelet Tour of Signal Processing*, Academic Press, 1999.
- [14] OSOWSKI S., *Metody i narzędzia eksploracji danych*, BTC, Legionowo, 2013.
- [15] LI H., JIANG T., *A class of edit kernels for SVMs to predict translation initiation sites in eukaryotic mRNAs*, J. of Computational Biology, vol. 12(6), 2005, pp. 702-718.
- [16] MCCULLAGH P., WANG H., ZHENG H., LIGHTBODY G., MCALLISTER G., *A comparison of supervised classification methods for auditory brainstem response determination*, Studies in Health Technology and Informatics, vol. 129, 2007, pp. 1289-1293.

M. SUCHOCKI, A. DOBROWOLSKI, E. MAJDA-ZDANCEWICZ,
K. TOMCZYKIEWICZ

Classification of Auditory Evoked Potentials based on the wavelet decomposition and SVM network

Abstract. For electrophysiological hearing assessment and diagnosis of brain stem lesions, the most often used are auditory brainstem evoked potentials of short latency. They are characterized by successively arranged maxima as a function of time, called waves. Morphology of the course, in particular, the timing and amplitude of each wave, allow a neurologist to make diagnose, what is not an easy task. A neurologist should be experienced, concentrated, and should have very good perception. In order to support his diagnostic process, the authors have developed an algorithm implementing the automated classification of auditory evoked potentials to the group of pathological and physiological cases, the sensitivity and specificity determined for an independent test group (of 50 cases) of respectively 84% and 88%.

Keywords: biomedical engineering, brainstem auditory evoked potentials, wavelet decomposition, support vector machine

DOI: 10.5604/12345865.1186236

