

Alina MOMOT
Politechnika Śląska, Instytut Informatyki

WAŻONE UŚREDNIANIE SYGNAŁU EKG WYKORZYSTUJĄCE ROZMYTY PODZIAŁ SYGNAŁU ORAZ WNIOSKOWANIE BAYESOWSKIE

Streszczenie. Akwizycja sygnału EKG wymaga tłumienia zakłóceń, co w przypadku sygnałów quasi-cyklicznych może być dokonywane przez ich uśrednianie. W warunkach rzeczywistych obserwuje się zmienność szumu z cyklu na cykl, co stanowi motywację dla stosowania metod ważonego uśredniania. W artykule proponuje się nową metodę uśredniania przy użyciu rozmytego podziału sygnału i bayesowskiego wnioskowania.

Słowa kluczowe: ważne uśrednianie, sygnał EKG

FUZZY WEIGHTED AVERAGING OF ECG SIGNAL USING FUZZY PARTITIONING OF THE SIGNAL AND BAYESIAN

Summary. Acquisition of ECG signals needs noise attenuation which, in case of quasi-cyclic signals, may be made by means of averaging. In reality the variability of noise power from cycle to cycle is observed, which constitutes a motivation for using methods of weighted averaging. This paper proposes a new weighted method incorporating fuzzy partitioning of the signal and Bayesian inference.

Keywords: weighted averaging, ECG signal

1. Wprowadzenie

Podstawowym celem elektrokardiografii jest diagnostyka pracy serca na podstawie obserwowalnych zjawisk elektrycznych. Zdolność włókien mięśni serca do spontanicznego przekazywania pobudzeń powoduje, że w skurczu biorą udział wszystkie zdolne do tego komórki mięśniowe. Systematyczna analiza rytmu EKG pozwala na dokładną ocenę zaburzeń tego

sygnału w stopniu wystarczającym do podjęcia bezpiecznego i skutecznego leczenia, a zapis tych zaburzeń i monitorowanie rytmu zatokowego może dostarczyć istotnych informacji diagnostycznych i przyczynić się do podjęcia właściwego leczenia długoterminowego [6]. Ze względu na brak możliwości „zbliżenia” procesu pomiarowego sygnału EKG do przedmiotu pomiaru, jakim jest serce, najczęściej stosuje się nieinwazyjną technikę akwizycji sygnału elektrycznego z powierzchni skóry, za pomocą elektrod rozmieszczonych na klatce piersiowej i kończynach. W związku z czym istnieje konieczność zaakceptowania dodatkowych źródeł sygnałów obocznych, rejestrowanych wspólnie z zapisem pracy serca. Zakłócenia obecne w zapisie EKG pochodzą zasadniczo z trzech źródeł, z których tylko jedno – aparatura pomiarowa – może być kontrolowane. Pozostałe źródła zakłóceń stanowią środowisko elektromagnetyczne, w którym odbywa się zapis oraz aktywność elektryczna innych organów. Źródła te, określane jako „przydźwięk sieci” oraz „zakłócenia mięśniowe”, mają najczęściej nieznaną i trudną do określenia charakterystykę [1].

Podstawowym celem analizy zapisów EKG jest wyodrębnianie w tym zapisie cech charakterystycznych dla określonych patologii. Często przy tym stosowane jest porównanie reprezentacji sygnału bieżącego ze wzorcem sygnału patologii. Powstaje zatem potrzeba tworzenia baz danych przechowujących wzorcowe zapisy zarówno sygnałów fizjologicznych, jak i patologicznych. Warto zauważyć przy tym, że w zapisie sygnału EKG przyjmowane są następujące założenia dotyczące dyskretyzacji sygnału ciągłego: stała precyzja amplitudowa oraz stały interwał czasowy próbkowania. Dla typowych systemów do realizacji badań elektrokardiograficznych akwizycja sygnału odbywa się z częstotliwością 1000 – 2000 Hz oraz z rozdzielczością 14 – 16 bitów. Zatem, oprócz wynikowych ciągów dyskretyzowanych danych w bazach przechowuje się również wartości częstotliwości próbkowania oraz precyzji dyskretyzacji amplitudy w postaci wyrażonej w jednostkach napięcia przypadających na najmniej znaczący bit.

Zaletą referencyjnych baz danych jest opis zapisanych tam sygnałów, zapewniający maksymalną wierność wartości i parametrów w stosunku do nieznanymi (z powodu występujących zakłóceń) wartości rzeczywistych. Europejska baza sygnałów syntetycznych CTS jest jedną ze standardowych baz danych, zawierającą zapisy elektrokardiograficzne wykorzystywane w celu testowania powstających na całym świecie algorytmów przetwarzania i analizy sygnałów EKG [12]. W niniejszym artykule przedstawiona zostanie nowa metoda uśredniania sygnału EKG, wykorzystująca rozmyty podziału sygnału i wnioskowanie bayesowskie, a weryfikacja użyteczności nowo powstałego algorytmu tłumienia zakłóceń obecnych w sygnale EKG zostanie dokonana właśnie przy użyciu sygnałów z bazy CTS.

Metody ważonego uśredniania, stosowane w celu zmniejszenia zakłóceń rejestrowanych wspólnie z zapisem pracy serca, wykorzystują quasi-cykliczny charakter sygnału EKG, polegający na występowaniu sygnałów o powtarzalnej morfologii w dziedzinie czasu. Uwzględ-

niając obserwowalną w praktyce zmienność poziomu zakłóceń, metody te opierają się na przydzielaniu najmniejszych wag tym cyklom, których poziom zakłóceń jest największy. Ocena poziomu zakłóceń w danym pobudzeniu jest zwykle szacowana na podstawie uśrednionego sygnału, co skutkuje potrzebą stosowania metod iteracyjnych [9, 11]. Oczywiście, tłumienie zakłóceń obecnych w sygnale EKG może być dokonywane również za pomocą innego typu metod wykorzystujących na przykład rozmytą maszynę, opartą na wektorach istotnych [10], analizę falkową [2], czy też analizę składowych głównych [7].

2. Metody uśredniania sygnału EKG

Niech w każdym cyklu $y_i(j)$ będzie sumą j -tej próbki deterministycznego sygnału $x(j)$, który jest taki sam w każdym pobudzeniu, i losowego szumu gaussowskiego $n_i(j)$ o średniej zero i wariancji stałej w każdym cyklu równej σ_i^2 . Zatem

$$y_i(j) = x(j) + n_i(j) \quad i = 1, 2, \dots, M \quad j = 1, 2, \dots, N, \quad (1)$$

gdzie i jest numerem cyklu, natomiast j jest numerem próbki w pojedynczym cyklu (wszystkie cykle mają tę samą długość N). Następujący wzór określa wartość j -tej próbki uśrednionego sygnału:

$$v(j) = \sum_{i=1}^M w_i y_i(j), \quad (2)$$

gdzie w_i jest wagą i -tego cyklu, a $v = [v(1), v(2), \dots, v(N)]^T$. Zazwyczaj dodatkowo zakłada się, że wagi są liczbami nieujemnymi, a ich suma wynosi 1, tak aby otrzymać estymator nieobciążony sygnału deterministycznego.

2.1. Tradycyjne uśrednianie arytmetyczne

W tradycyjnym uśrednianiu arytmetycznym przyjmuje się, że wszystkie wagi są sobie równe, co, uwzględniając warunek mówiący, że ich suma ma wynosić 1, prowadzi do wartości M^{-1} . Jeśli wariancja szumu jest taka sama w każdym cyklu, tj. $\sigma_i^2 = \sigma^2 \quad \forall i = 1, 2, \dots, M$, to w ten sposób określone wagi są optymalne w sensie minimalizowania błędu średniokwadratowego pomiędzy sygnałem uśrednionym v oraz sygnałem deterministycznym x , zakładając nadal gaussowski charakter szumu. Natomiast w przypadku, gdy zakłócenia nie charakteryzują się gaussowskim rozkładem, uśredniony sygnał opisany wzorem (2) nie stanowi optymalnego rozwiązania, ale jest optymalnym liniowym estymatorem sygnału x [8].

W przypadku uwzględniającym możliwość różnych wariancji szumu w poszczególnych cyklach klasyczna procedura dla wyznaczenia średniej ważonej prowadzi do wartości wag, które są proporcjonalne do odwrotności tych wariancji [4]:

$$w_i = \sigma_i^{-2} \left(\sum_{m=1}^M \sigma_m^{-2} \right)^{-1}. \quad (3)$$

W takim przypadku wariancja uśrednionego sygnału jest dana wzorem:

$$\text{Var } v(j) = \left(\sum_{i=1}^M \sigma_i^{-2} \right)^{-1}, \quad (4)$$

podczas gdy wariancja sygnału uśrednionego w sposób arytmetyczny wynosi

$$\text{Var } v(j) = \sum_{i=1}^M \frac{\sigma_i^2}{M^2}. \quad (5)$$

Jeśli wariancja szumu jest taka sama w każdym cyklu, tj. $\sigma_i^2 = \sigma^2 \quad \forall i = 1, 2, \dots, M$, prowadzi to oczywiście do uzyskania takiej samej wartości wariancji równej

$$\text{Var } v(j) = \frac{\sigma^2}{M}. \quad (6)$$

2.2. Wykorzystanie wnioskowania bayesowskiego do uśredniania

W warunkach rzeczywistych podczas akwizycji sygnału EKG obserwuje się zmienność poziomu szumu z cyklu na cykl, co stanowi motywację dla stosowania metod ważonego uśredniania. Zakładając, że wariancje szumu σ_i^2 w każdym cyklu $i = 1, 2, \dots, M$ są znane, wagi mogą być wyznaczone wprost stosując wzór (3), jednak wyznaczenie tych wariancji jest niemożliwe. Tak więc wariancja szumu jest albo estymowana na podstawie pobranego sygnału EKG, albo stosowane są inne metody wyznaczania optymalnych wag bez bezpośredniego estymowania tych wariancji. Poniżej opisana zostanie bayesowska metoda wyznaczania sygnału uśrednionego zakładająca, że sygnał deterministyczny jest zmienną losową, wykorzystująca jej rozkład a posteriori warunkowany danymi wejściowymi, czyli sygnałem z zakłóceniami, gdzie wariancja szumu jest estymowana za pomocą metody momentów.

Zakładając, że dany jest zbiór $y = \{y_i(j)\}$, gdzie i jest indeksem cyklu $i = 1, 2, \dots, M$, natomiast j jest indeksem próbki w cyklu $j = 1, 2, \dots, N$. Niech $y_i(j) = x(j) + n_i(j)$, gdzie $n_i(j)$ jest szumem gaussowskim o średniej zero i wariancji w i -tym cyklu równej σ_i^2 , natomiast sygnał x ma rozkład gaussowski o średniej zero i przekątnej macierzy kowariancji $B = \text{diag}(\beta_1^{-1}, \beta_2^{-1}, \dots, \beta_N^{-1})$, gdzie $\beta_1^{-1} = \beta_2^{-1} = \dots = \beta_N^{-1} = \beta$. Zatem, na podstawie twierdzenia Bayesa [5] rozkład a posteriori dla sygnału x jest dany wzorem

$$p(x|y, \beta) = \frac{p(y|\beta, x)p(x|\beta)}{\int p(y|\beta, x)p(x|\beta)dx}, \quad (7)$$

co można przekształcić do prostszej postaci w sposób analityczny, otrzymując rozkład normalny o wektorze średnich $m = [m(1), m(2), \dots, m(N)]^T$:

$$m(j) = \frac{\sum_{i=1}^M \sigma_i^{-2} y_i(j)}{\beta + \sum_{i=1}^M \sigma_i^{-2}}, \quad \text{dla } j = 1, 2, \dots, N \quad (8)$$

m oraz macierzy kowariancji $C = c I_N$, gdzie

$$c = \left(\beta + \sum_{i=1}^M \sigma_i^{-2} \right)^{-1} \quad (9)$$

natomiast I_N jest macierzą jednostkową wymiaru N .

Jako oszacowanie wektora x można zatem przyjąć wektor średnich m rozkładu prawdopodobieństwa $p(x|y, \beta)$, nieznanne zaś parametry występujące we wzorze (8), czyli odwrotności wariancji rozkładu a priori sygnału deterministycznego – β oraz wariancji szumu – σ_i^{-2} , można estymować metodą momentów [3], co prowadzi do otrzymania

$$\beta = \frac{N}{\sum_{j=1}^N (x(j))^2} \quad (10)$$

oraz

$$\sigma_i^{-2} = \frac{N}{\sum_{j=1}^N (y_i(j) - x(j))^2} \quad \text{dla } i = 1, 2, \dots, M. \quad (11)$$

2.3. Opis proponowanego algorytmu

Iteracyjny algorytm opisanej w poprzedniej sekcji bayesowskiej metody ważonego uśredniania można przedstawić następująco:

1. Ustalić $v^{(0)} \in R^N$ (np. zastosować tradycyjne uśrednianie arytmetyczne) i ustawić indeks iteracji $k = 1$.
2. Wyznaczyć parametr $\beta^{(k)}$ przy użyciu wzoru (10) oraz $(\sigma_i^{-2})^{(k)}$ przy użyciu wzoru (11) dla $i = 1, 2, \dots, M$, przyjmując, że $x = v^{(k-1)}$.
3. Wyznaczyć uśredniony sygnał k -tej iteracji $v^{(k)}$ przy użyciu wzoru (8) oraz wyznaczonych w poprzednim etapie wartości $\beta^{(k)}$ oraz $(\sigma_i^{-2})^{(k)}$, przyjmując, że $v^{(k)} = m$.
4. Jeżeli $\|v^{(k)} - v^{(k-1)}\| > \varepsilon$, to $k \leftarrow k + 1$ i iść do etapu 2.

Warto przy tym zwrócić uwagę na fakt, że opisany wyżej algorytm nie wymaga ustalania żadnych dodatkowych parametrów, w odróżnieniu od podobnych algorytmów bayesowskiego ważonego uśredniania opisanych w pracy [11].

Na podstawie tak określonego algorytmu można dokonać jego rozszerzenia, polegającego na rozmytym podziale sygnału podlegającemu uśrednianiu. Proponuje się podział na K klas przez nałożenie gaussowskich funkcji przynależności

$$f_{a,b}(x) = \exp\left\{-\left(\frac{x-a}{b}\right)^2\right\}, \quad (12)$$

o różnych parametrach środka a równomiernie przesuwających się wzdłuż numerów próbek od 1 do N , na sygnał wejściowy $y_i = [y_i(1), y_i(2), \dots, y_i(N)]^T$ dla każdego $i = 1, 2, \dots, M$, natomiast parametr b byłby równy $1/4$ szerokości okna, tzn.

$$a^k = \frac{1}{2} \frac{N}{K} + (k-1) \frac{N}{K} = \left(k - \frac{1}{2}\right) \frac{N}{K}, \quad \text{dla } k = 1, 2, \dots, K \quad (13)$$

$$b = \frac{1}{4} \frac{N}{K}. \quad (14)$$

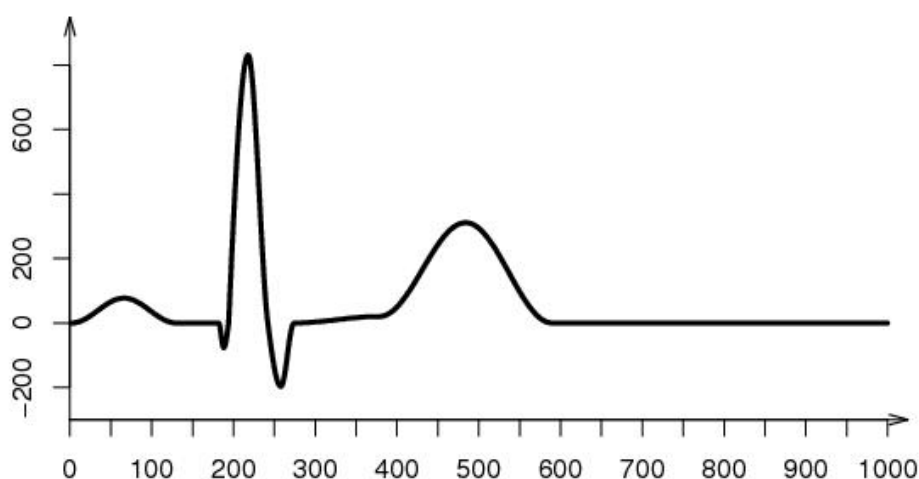
Zatem, otrzymuje się nowe zestawy danych wejściowych y_i^k dla $k = 1, 2, \dots, K$, gdzie

$$y_i^k(j) = y_i(j) \cdot \exp\left\{-\left(\frac{j-a^k}{b}\right)^2\right\}, \quad (15)$$

które będą niezależnie w każdej z K klas uśredniane, a następnie wyniki będą sumowane, dając jeden uśredniony sygnał.

3. Eksperymenty numeryczne

Poniżej zostaną opisane wyniki eksperymentów numerycznych weryfikujących skuteczność opisywanej metody w zestawieniu z tradycyjnym uśrednianiem arytmetycznym (TUA) oraz istniejącymi metodami ważonego uśredniania, takimi jak metoda WACFM opisana w pracy [9], bayesowskie metody EBWA.1 i EBWA.3 opisane w pracy [11]. Jako składnik deterministyczny wybrano sygnał elektrokardiograficzny ANE20000 (zaprezentowany na rys. 1) pochodzący z bazy danych CTS [13]. Eksperymenty były przeprowadzone w ogólnodostępnym środowisku do obliczeń statystycznych i grafiki **R** w wersji 2.4.0 (<http://www.r-project.org>). Dla wyznaczonych uśrednionych sygnałów skuteczność redukcji zakłóceń porównywano wykorzystując maksymalną wartość bezwzględną różnicy między sygnałem oryginalnym a uśrednionym (MAX) oraz pierwiastek błędu średniokwadratowego między sygnałem oryginalnym a uśrednionym (RMSE).



Rys. 1. Sygnał ANE20000

Fig. 1. ANE20000 signal

W pierwszej fazie eksperymentu sygnał ANE 20000 został 100-krotnie powielony i zakłócony szumem gaussowskim o odpowiednio $0.1s$, $0.5s$, $1s$, $2s$ (gdzie s jest próbkowym odchyleniem standardowym sygnału użytecznego) w każdym z czterech 25-elementowych ciągów powielonego sygnału ANE20000, a wyniki eksperymentu przedstawia tabela 1.

Prezentowane rezultaty otrzymane były w wyniku zastosowania znanych metod, nowej metody bayesowskiego ważonego uśredniania (BWU) oraz rozszerzenia tej metody wykorzystującej rozmyty podział sygnału (RBWU). W algorytmach ważonego uśredniania wartością początkową $v^{(0)}$ było tradycyjne uśrednienie arytmetyczne, a jako ε przyjęto wartość 10^{-10} . Dla tych parametrów liczba iteracji dla wszystkich metod nie przekraczała kilkunastu.

W drugiej fazie eksperymentu sygnał ANE20000 został 100-krotnie powielony i zakłócony szumem o rozkładzie Cauchy'ego, który może służyć jako model zakłóceń impulsowych. Funkcja gęstości prawdopodobieństwa tego rozkładu wyraża się wzorem:

$$f(x) = \frac{1}{\pi\sigma} \left(1 + \left(\frac{x-\mu}{\sigma} \right)^2 \right)^{-1} \quad (16)$$

gdzie μ jest parametrem położenia, natomiast σ jest parametrem skali.

Tabela 1

Wyniki eksperymentów dla gaussowskiego szumu

Metoda uśredniania	Parametry	RMSE	MAX
TUA		12.88784	44.78031
WACFM	m = 2	2.628207	8.974255
EBWA.1	p = 1	2.551379	9.308891
	p = 2	2.554577	9.329751
	p = 3	2.555229	9.333002
EBWA.3	p = 2	2.555382	9.339712
	p = 3	2.555559	9.342102
BWU		2.556326	9.342494
RBWU	K = 2	2.552506	9.480585
	K = 3	2.049810	9.608289
	K = 4	2.184384	9.506229
	K = 5	1.975438	9.376102

Parametry te są zwykle używane w miejsce wartości oczekiwanej i odchylenia standardowego z uwagi na brak dwóch pierwszych momentów w rozkładzie Cauchy'ego. W tej fazie eksperymentu parametr μ wynosił zero, natomiast σ był równy $0.05s$ (gdzie s jest, podobnie jak w przypadku pierwszej fazy, próbkowym odchyleniem standardowym sygnału użytecznego). Wyniki tej fazy eksperymentu przedstawia tabela 2.

Wyniki przeprowadzonych eksperymentów numerycznych świadczą o tym, że wszystkie metody ważonego uśredniania, jakkolwiek bardziej złożone obliczeniowo w stosunku do tradycyjnego uśredniania arytmetycznego, charakteryzują się jednak znacznie wyższą skutecznością redukcji poziomu zakłóceń w przypadku zmiennej amplitudy szumu gaussowskiego, jak również w przypadku zakłóceń szumem Cauchy'ego. Zaprezentowana w artykule nowa metoda bayesowskiego ważonego uśredniania (BWU) daje wyniki podobne do innych metod bayesowskich (EBWA.1 oraz EBWA.3) traktowanych jako metody referencyjne w przeprowadzonych eksperymentach numerycznych. Jednak rozszerzenie tej metody, polegające na dodatkowym zastosowaniu rozmytego podziału sygnału, prowadzi do uzyskania znacznie lepszych rezultatów. Wraz ze zwiększającą się liczbą przedziałów można zaobserwować zmniejszające się wartości pierwiastków błędów średniokwadratowych, a także mak-

symalnych wartości bezwzględnych różnic między sygnałem uśrednionym a sygnałem oryginalnym.

Tabela 2

Wyniki eksperymentów dla szumu Cauchy'ego

Metoda uśredniania	Parametry	RMSE	MAX
TUA		747.6133	19343.1406
WACFM	$m = 2$	81.07113	1773.52767
EBWA.1	$p = 1$	18.13755	121.13051
	$p = 2$	18.80865	121.72559
	$p = 3$	18.96972	122.01273
EBWA.3	$p = 2$	19.33862	125.32701
	$p = 3$	19.53295	126.75974
BWU		19.59817	126.52550
RBWU	$K = 2$	13.60974	81.50904
	$K = 3$	8.701535	46.453579
	$K = 4$	8.708482	45.301270
	$K = 5$	6.810976	32.441117

Wartościowe wydaje się w przyszłości zbadanie zachowania się zaproponowanego rozszerzenia opisanej metody baysowskiego ważonego uśredniania do innych metod uśredniających, a także przebadanie skuteczności takich rozszerzeń w przypadku zastosowania innych niż gaussowska funkcji przynależności lub też innego typu podziału niż podział równomierny.

Praca naukowa sfinansowana ze środków na naukę w latach 2007-2009 jako projekt badawczy nr N N518 1200 33.

BIBLIOGRAFIA

1. Augustyniak P.: Przetwarzanie sygnałów elektrodiagnostycznych. Uczelniane Wydawnictwa Naukowo-Dydaktyczne AGH, Kraków 2001.
2. Augustyniak P.: Adaptive wavelet discrimination of muscular noise in the ECG. *Computers in Cardiology*, Vol. 33, 2006, s. 481÷484.
3. Bartoszewicz J.: Wykłady ze statystyki matematycznej. PWN, Warszawa 1996.
4. Brandt S.: *Statistical and Computational Methods in Data Analysis*. Springer Verlag, 1997.
5. Carlin B., Louis T.: *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman & Hall, New York 1996.
6. Dąbrowska B., Dąbrowski A.: *Podręcznik elektrokardiografii*. Wydawnictwo Lekarskie PZWL, Warszawa 1999.

7. Kotas M.: Application of projection pursuit based robust principal component analysis to ECG enhancement, *Biomedical Signal Processing and Control*, Vol. 1, No. 4, 2006, s. 289÷298.
8. Łęski J.: Zastosowanie metody uśredniania sygnału w dziedzinie czasu i filtracji Kalmana do tłumienia zakłóceń sygnału EKG. Rozprawa doktorska. Politechnika Śląska, Gliwice 1989.
9. Łęski J.: Robust Weighted Averaging. *IEEE Transactions on Biomedical Engineering*, Vol. 49, No. 8, 2002, s. 796÷804.
10. Momot A., Momot M., Łęski J.: The Fuzzy Relevance Vector Machine and its Application to Noise Reduction in ECG Signal. *Journal of Medical Informatics and Technologies*, Vol. 9, 2005, s. 99÷106.
11. Momot A., Momot M., Łęski J.: Bayesian and empirical Bayesian approach to weighted averaging of ECG signal. *Bulletin of the Polish Academy of Science. Technical Science*, Vol. 55, No. 4, 2007, s. 341÷350.
12. European Conformance Testing Services for Computerized Electrocardiography, CTS ECG Test Atlas. BIOSIGNA, (www.biosigna.de/CTS.pdf).
13. International Electrotechnical Commission Standard 60601-3-2, 1999.

Recenzent: Prof. dr hab. inż. Andrzej Świerniak

Wpłynęło do Redakcji 12 stycznia 2009 r.

Abstract

Averaging signals in time domain is one of the main methods of noise attenuation in biomedical signal processing in case of systems producing repetitive patterns such as ECG acquisition systems. This paper presents new method of Bayesian weighted averaging of ECG signal with extension using fuzzy partitioning of the signal. Performance of the presented methods is experimentally evaluated and compared with the traditional averaging by using arithmetic mean and well known weighted averaging methods based on criterion function minimization WACFM or based on empirical Bayesian inference EBWA.1 and EBWA.3. The presented performance comparison of all described methods uses a synthetic signal ANE20000 from CTS database. A series of 100 ECG cycles was generated with the same deterministic component and zero-mean white Gaussian noise with different standard deviations with constant amplitude of noise during each cycle or Cauchy noise which was treated

as simulated impulse noise. The results of experiments show supremacy of the new method using Bayesian inference and fuzzy partitioning of the signal with regard to both type of noise especially in case of increasing fuzzy intervals number.

Adres

Alina MOMOT: Politechnika Śląska, Instytut Informatyki, ul. Akademicka 16,
44-100 Gliwice, Polska, alina.momot@polsl.pl.