



Akademie věd České republiky
Ústav teorie informace a automatizace

Academy of Sciences of the Czech Republic
Institute of Information Theory and Automation

RESEARCH REPORT

M. ZIMA, P. TICHAVSKÝ A V. KRAJČA:

Odstraňování artefaktů v EEG datech III

No. 2259

Září 2009

ÚTIA AV ČR, P. O. Box 18, 18208 Prague, Czech Republic
Telex: 122018 atom c, Fax: (+42) (2) 688 4903
E-mail: utia@utia.cas.cz

This report constitutes an unrefereed manuscript which is intended to be submitted for publication. Any opinions and conclusions expressed in this report are those of the author(s) and do not necessarily represent the views of the Institute.

Úvod

Data z elektroencefalografu (EEG) bývají často poškozena nežádoucími signály nejrůznějšího druhu, tzv. artefakty, které ztěžují interpretaci vlastní mozkové aktivity. Tato práce si klade za cíl navrhnout algoritmus, který by pomohl k odstraňování těchto artefaktů. Hlavní důraz klademe na artefakty, které trvají relativně krátkou dobu vzhledem ke zbytku signálu.

Naše práce se skládá ze čtyř částí. První část se zabývá aplikací metody analýzy nezávislých komponent na odstraňování artefaktů z krátkého EEG záznamu (~20 sekund). V druhé části jsou použity postupy z první části na záznam libovolné délky. Třetí část se zabývá srovnáním algoritmů pro rozklad signálu na nezávislé komponenty. V prvních třech částech jsou pro možnost co nejobjektivnějšího posouzení úspěšnosti navrhovaného algoritmu zkoumány výsledky na datech, do kterých modely námi zkoumaných artefaktů sami vkládáme. Ve čtvrté části je pak prezentována funkčnost na datech se skutečnými artefakty.

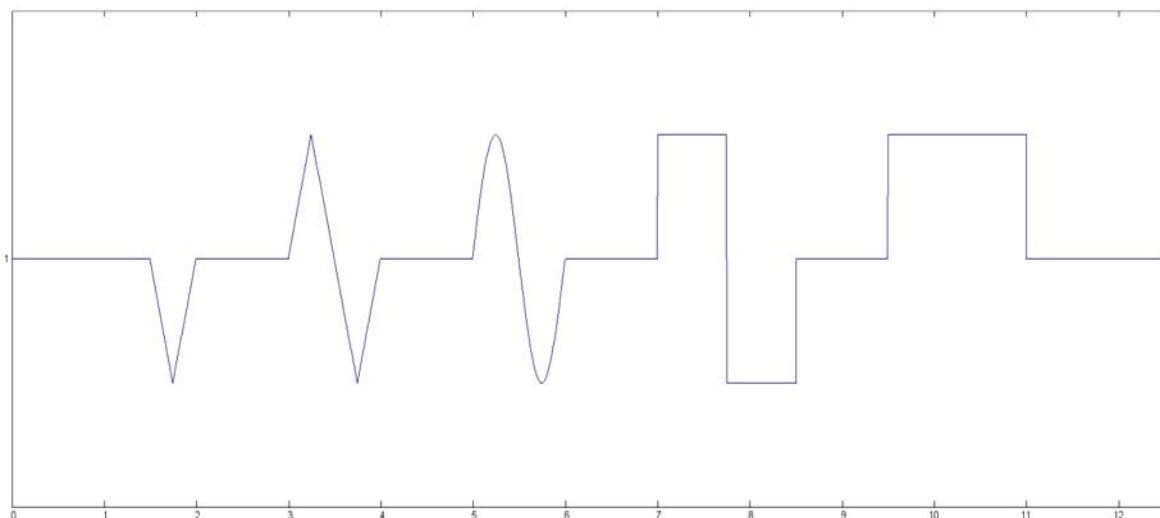
Odstraňování artefaktů z krátkého záznamu

Zpracovávaná data

Data, nad kterými jsme prováděli testy, jsme získali z Ústavu péče o matku a dítě v Praze 4 Podolí. Jedná se o novorozenecké EEG, což je patrné z počtu kanálů, kterých je v tomto případě 8. Signál z EEG u dospělých obsahuje kanálů 19 a víc. Veškeré závěry, které v této práci činíme, jsou platné pro libovolný počet kanálů. V případě většího počtu kanálů se úloha odstraňování artefaktů zjednodušuje, neboť máme k dispozici více informací k rekonstrukci poškozených dat.

Modely artefaktů

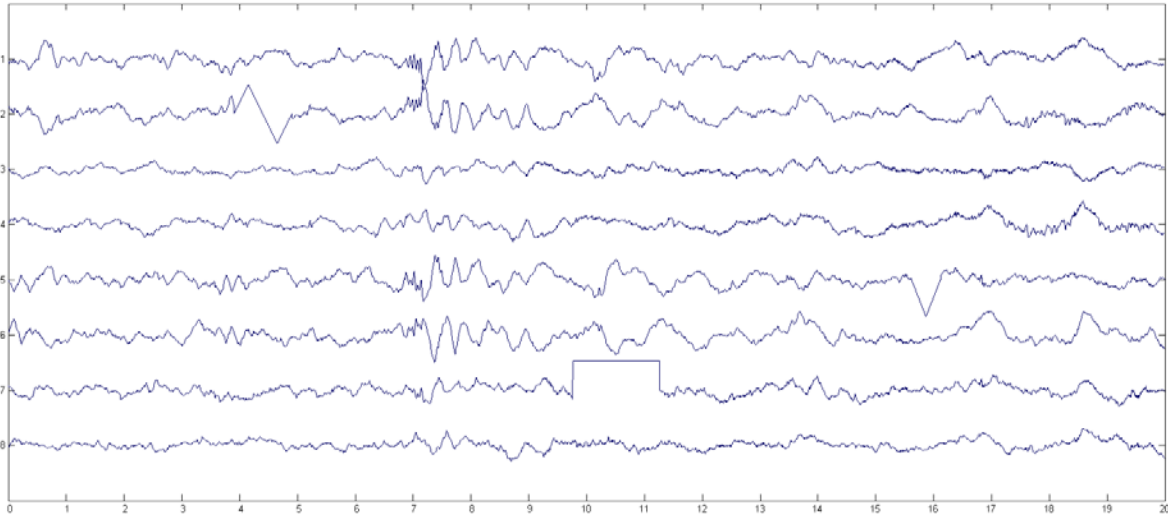
Pro účely simulace artefaktů jsme zvolili pět modelů, které jsou zobrazeny na obr. 1. Čísla v jeho spodní části označují čas v sekundách.



Obr. 1 – modely artefaktů

Popis úlohy

Na obr. 2 jsou data, do kterých jsme uměle vložili 3 artefakty. Naší úlohou je tyto nežádoucí poruchy odstranit.



Obr. 2 – data se 3 artefakty (2. kanál – 4. sekunda; 5. kanál – 10. sekunda 7. kanál – 16. sekunda)

Analýza nezávislých komponent

Veškeré námi navržené algoritmy na odstraňování artefaktů se opírají o metodu analýzy nezávislých komponent. Při této metodě je nutné původní signál sestávající z n -tice kanálů (komponent) transformovat tak, aby komponenty výsledného signálu byly co nejvíce nezávislé.

Hlavní motivací pro použití metody analýzy nezávislých komponent je předpoklad, že artefakt je část signálu, která se od zbytku nějakým zásadním způsobem liší. Bude proto po separaci původního signálu do nezávislých komponent odfiltrován do samostatné komponenty. Pokud k tomu dojde, můžeme tento nežádoucí signál nahradit signálem nulovým a inverzní transformací přejít k původní n -tici signálů, která již v ideálním případě artefakt obsahovat nebude.

Největší pozornost je věnována lineárnímu modelu

$$\mathbf{X} = \mathbf{A}\mathbf{S}$$

kde

$$\mathbf{X} = \begin{pmatrix} x_{11} & \cdots & x_{1N} \\ \vdots & \ddots & \vdots \\ x_{d1} & \cdots & x_{dN} \end{pmatrix}$$

značí matici N realizací d naměřených hodnot. Dále \mathbf{S} značí matici dat, které tvoří nezávislé komponenty. Konečně \mathbf{A} je neznámá $d \times d$ regulární mixovací matice.

K určení matice \mathbf{A} , přesněji k určení matice k ní inverzní, máme k dispozici několik následujících metod:

EFICA (Efficient Fast ICA)– asymptoticky eficientí algoritmus, navržený Koldovským, Tichavským a Ojou [6], který je vylepšením populárního algoritmu FastICA [1]. Je založen na optimalizaci vhodné nelineární účelové funkce a je zobecněním maximalizace kurtosy signálů.

WASOBI (Weigh-Adjusted Second-Order Blind Identification)– metoda separace opírající se o spektrální diverzitu (různost spekter) separovaných signálů [7]. Využívá pouze statistik druhého řádu a je to zdokonalená verze populárního algoritmu SOBI [2]. Separace je statisticky eficientí (dosahuje Rao-Cramerovy hranice), pokud separované procesy jsou Gaussovsky rozložené autoregresní procesy daného řádu. Tento předpoklad ovšem nemusí být vůbec splněn pro úspěšnou separaci.

BGSEP (Block Gaussian Separation)– metoda založená na nestacionaritě separovaných signálů. Provádí se pomocí rozdělení signálu na segmenty, výpočtu kovariančních matic na jednotlivých segmentech, a jejich přibližnou vzájemnou diagonalizací těchto matic [10].

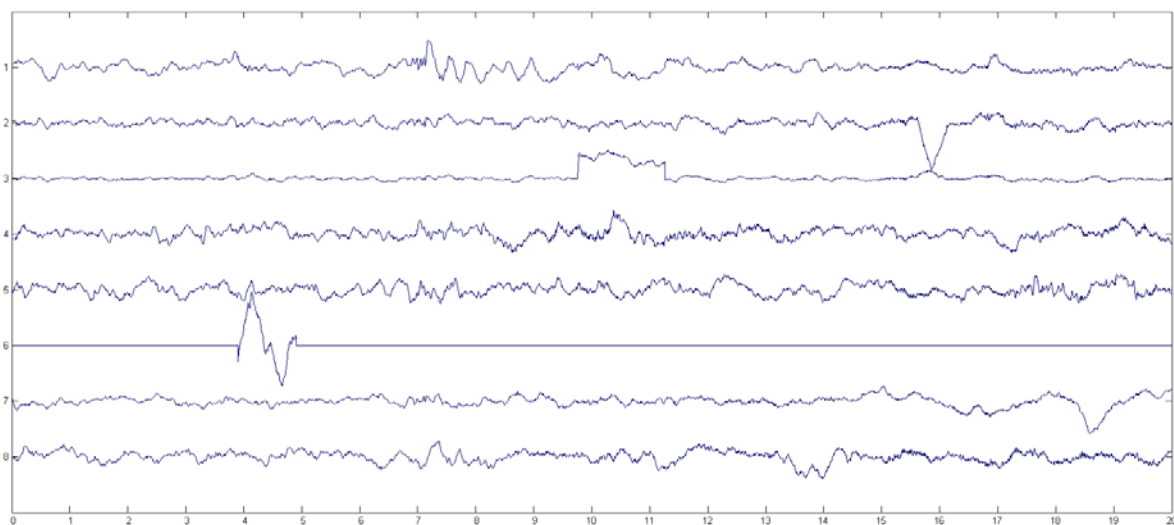
BARBI (Block AutoRegressive Blind Identification)– zobecnění algoritmů WASOBI a BARBI. Signál je rozdělen na bloky a v každém bloku je každý separovaný signál modelován jako autoregresní proces určitého řádu. Má dva parametry, počet bloků a řád autoregrese. BARBI(10,0) je ekvivalentní BGSEP s 10 bloky a BARBI (1,10) je ekvivalentní WASOBI s řádem autoregrese 10.

Srovnání jednotlivých metod je věnována třetí část naší práce.

Pokud matici A máme k dispozici, můžeme od naměřeného signálu přejít k jeho nezávislým komponentám vynásobením původních dat touto maticí. Stejně jednoduchý je i inverzní přechod.

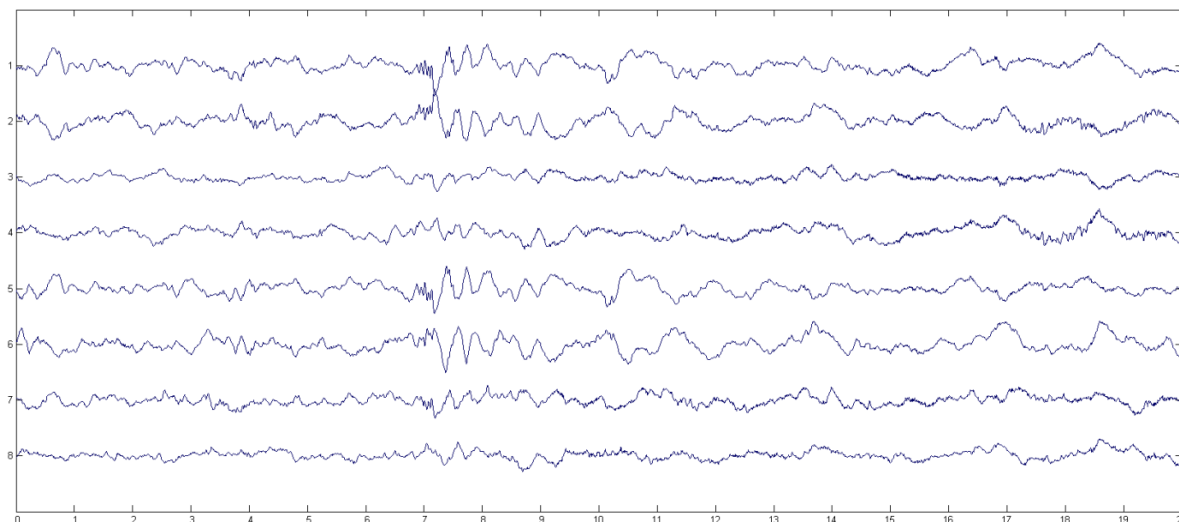
Aplikace analýzy nezávislých komponent při odstraňování artefaktů z krátkého záznamu

Nezávislé komponenty signálu z obr. 2 jsou na obr. 3.



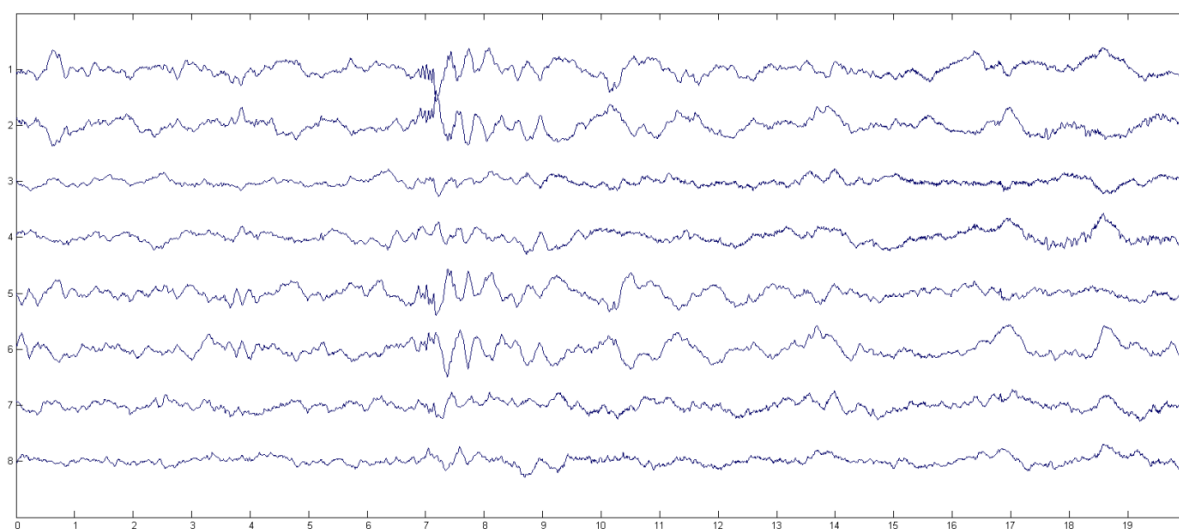
obr. 3 – nezávislé komponenty

Po odstranění 2,3 a 6 komponenty a přejití k původnímu signálu dostaneme signál na obr. 4.

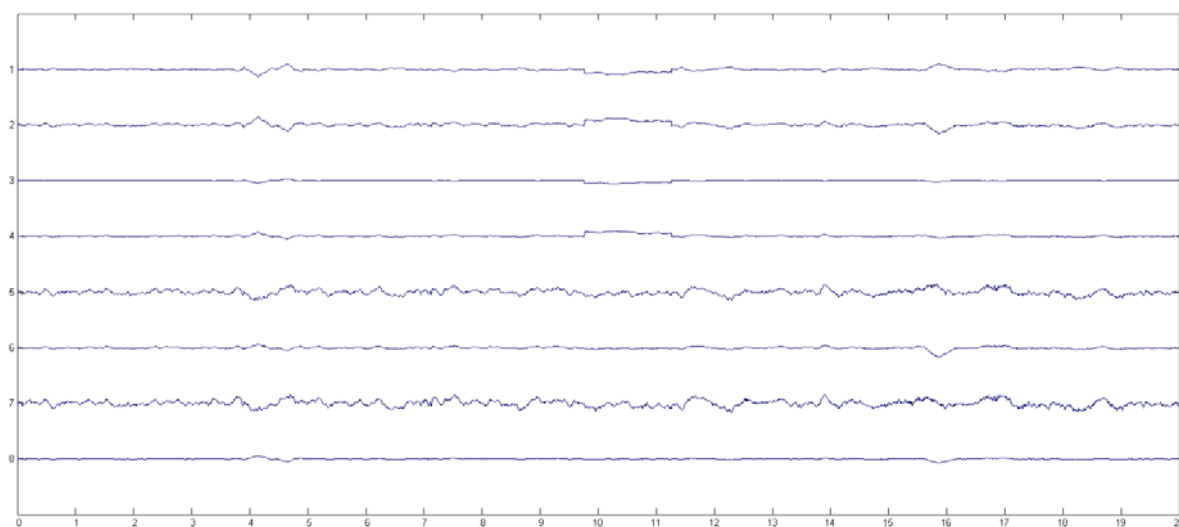


obr. 4 – signál po odstranění komponent obsahujících artefakt

Na obr. 5 je původní signál před přidáním artefaktů. Pro snazší srovnání původního signálu a signálu po vyčištění ukazuje obr. 6 jejich rozdíl.



obr. 5 – původní signál



obr. 6 – rozdíl původního signálu a signálu po vyčištění

Srovnáním obr. 4 a 5, popřípadě z obr. 6 vidíme, že došlo k odstranění námi přidaných artefaktů. Při odstraňování artefaktů dále došlo k drobné modulaci výsledného signálu oproti původnímu. Patrné je to zejména z obr. 6 v 5. a 7. kanále. Nicméně tato modulace je v jednotlivých kanálech rozprostřena rovnoměrně a nemá tedy na posouzení celkového charakteru signálu vliv.

Odstraňování artefaktů z libovolného záznamu

Dělení na segmenty

Úloha, kterou se chceme v této práci zabývat je automatické odstraňování artefaktů z libovolně dlouhého záznamu. Prvním krokem tedy musí být rozdělení původního záznamu na segmenty, které budou dále analyzovány.

Segmenty by neměly být příliš dlouhé, aby sami neobsahovaly větší počet artefaktů a nezvyšovaly tak počet komponent, které budou při zpětné transformaci odstraněny. Zároveň však musí obsahovat dostatečné množství dat, aby algoritmus pro separaci nezávislých komponent dobře fungoval. Otázce optimální délky segmentu se dále věnujeme při aplikaci algoritmu na konkrétních datech.

Návrh dělit signál na segmenty pevně zvolené délky má jistě několik nevýhod plynoucích z nereflexování skutečného rozmístění artefaktů v analyzovaném signálu. Mohlo by se jevit výhodnější měnit délku segmentu dynamicky v závislosti na počtu artefaktů v dané části signálu. Bylo by tedy potřeba řešit optimalizační úlohu vzhledem k délce segmentu. Její vyřešení by však nemuselo nutně vést ke zlepšení výsledků, spíše by vedlo k větší časové náročnosti celého algoritmu. Lepším řešením se nám proto jeví volit kratší, leč pevně dlouhé segmenty.

Druhou námitkou proti dělení signálu na segmenty pevné délky je, že může docházet k rozdělení artefaktů do dvou sousedních segmentů, čímž se znemožňuje, nebo alespoň komplikuje jeho odstranění. K minimalizaci tohoto jevu původní signál analyzujeme ve třech nezávislých chodech, vždy s posunutými začátky segmentů. Výsledný signál poté po částech komponujeme z těchto dílčích výsledků, každý kanál zvlášť. Délka těchto částí může být obecně různá od délky segmentů. Pokud se při skládání signálů odhalí, že určitá část jednoho ze signálů se výrazně odlišuje od odpovídajících částí zbylých dvou, je tomu tak pravděpodobně proto, že se v dané části signálu neodstranila část artefaktu, zatímco v ostatních případech ano. Pokud se toto zjistí, ty z trojice, které mají významně větší amplitudu, se při rekonstrukci dané části signálu nepoužijí.

Pro účely analýzy je užitečné zvolit nějaké kritérium, kterým určíme rozdílnost dvou signálů v jednom kanále. Pro naše potřeby, vzhledem k tomu, že máme k dispozici pouze body, které danou křivku definují, jsme zvolili součet druhých mocnin rozdílu funkčních hodnot, tedy

$$\|x - y\|^2 = \sum_{i=1}^N (x_i - y_i)^2 \quad (1)$$

kde $x=(x_1, \dots, x_N)$ a $y=(y_1, \dots, y_N)$ jsou signály v jedné části nebo obecně v jednom kanálu pozorovaného úseku dat.

S takto definovanou metrikou můžeme podrobněji vysvětlit, jak při rekonstrukci signálu v jedné části z trojice rekonstrukcí určujeme, že se jedna z nich výrazně odlišuje od zbylých dvou.

Označme tyto rekonstrukce (ve formě vektorů) písmeny a, b, c . Necht' jejich vzdálenosti jsou $\rho_1 = \|a - b\|, \rho_2 = \|b - c\|, \rho_3 = \|a - c\|$ a necht' \tilde{x} je odmocnina střední kvadratické hodnoty signálu na celých datech (i vně uvažované části). Pokud $\max \rho_i > 2 * \min \rho_i$ a zároveň $2 * \tilde{x} > \max \rho_i$, potom řekneme, že se jedna z rekonstrukcí významně odlišuje od zbylých dvou.

Není samozřejmě jasné, po jak velkých částech provádět rekonstrukci výsledného signálu z dílčí trojice. V implementaci algoritmu jsme zvolili délku odpovídající 1 sekundě, čili jakousi charakteristickou časovou jednotku měření.

Odstraňování artefaktů uvnitř segmentů

Pokud se nám podaří odstranit problémy s artefakty na rozhraní jednotlivých segmentů, zbývá odstranit artefakty uvnitř. K tomu nám poslouží metoda analýzy nezávislých komponent popsaná v prvním oddílu naší práce. Tu nyní automatizujeme.

Vzhledem k třídě artefaktů, kterými se zabýváme, by měla komponenta obsahující artefakt mít převážně nulový průběh a aktivitu s výraznou amplitudou vykazovat jen v relativně krátkém časovém úseku. Výše uvedenou vlastnost je pro potřeby automatické detekce nutno nějak kvantifikovat. Protože artefakt vykazuje výraznější nenulové chování jen v krátkém časovém úseku, bude medián z absolutních hodnot signálu velmi blízko nule. Naproti tomu střední hodnota z absolutních hodnot bude díky výrazné amplitudě v nenulové části signálu relativně k hodnotě mediánu vysoká. Poměr střední hodnoty a mediánu z absolutních hodnot by tak mohl poměrně dobře charakterizovat přítomnost artefaktu v komponentě, neboť pro signál neobsahující artefakt tento poměr bude velmi blízko jedné.

Je dobré si uvědomit, že k dokonalému odfiltrování artefaktů do samostatných komponent v praxi nikdy nedojde. Vždy bude komponenta s artefaktem částečně modulována ostatním signálem. Pro účely identifikace je tak výhodné zesílit vlastnosti, podle kterých signál posuzujeme. K tomuto účelu neohodnocujeme signál rozložený do nezávislých komponent rovnou. Napřed každou z komponent vydělí tři čtvrtě násobkem maximální hodnoty a funkční hodnoty umocníme na desátou. Definujeme tedy

$$(2) \quad y_i^{(j)} = \left(\frac{x_i^{(j)}}{0.75 * \max(x^{(j)})} \right)^{10}$$

kde $x^{(j)} = (x_1^{(j)}, \dots, x_N^{(j)})$ je j -tá nezávislá komponenta a $y^{(j)} = (y_1^{(j)}, \dots, y_N^{(j)})$ odpovídající předzpracovaný signál určený k další analýze.

Anomalitu komponenty definujeme jako

$$(3) \quad f(y^{(j)}) = \log \left(\frac{\sum_{i=1}^N y_i^{(j)}}{\text{median}(y^{(j)})} \right).$$

Tato veličina charakterizuje do jaké míry tuto komponentu považujeme v datech za nežádoucí (artefakt).

Otázkou zůstává, pro jak vysoké hodnoty anomaly označíme ohodnocovaný signál za komponentu obsahující artefakt. Mez, při jejímž překročení komponentu určíme jako komponentu obsahující artefakt, ponecháváme jako volný parametr, který je potřeba odhadnout experimentálně. K tomuto problému se vracíme při aplikaci algoritmu na konkrétní data. Obecně však zřejmě platí, že když budeme mez snižovat, bude se počet komponent označených za komponenty obsahující artefakt zvyšovat.

Implementace algoritmu

1. Nastavení parametrů
 - a. Mezní hodnota pro anomalitu
 - b. Maximální počet odstraňovaných komponent
 - c. Metoda pro rozklad signálu na nezávislé komponenty
 - d. Délka segmentu pro analýzu signálu
 - e. Délka částí pro rekonstrukci
2. Získání dílčích rekonstrukcí s částečně odstraněnými artefakty
 - a. Rozdělení vstupního signálu na segmenty zadané délky a v každém segmentu
 - I. Zvolenou metodou určíme mixovací matici a signál v daném segmentu rozložíme na nezávislé komponenty
 - II. Určíme anomalitu jednotlivých komponent
 - III. Komponenty, jejichž anomalita je vyšší než zadaná mez, nahradíme nulovým signálem. Pokud by se mělo odstranit komponent více než maximální počet, odstraní se pouze ty s nejvyšší anomalitou.
 - IV. Inverzní transformací přejdeme k původnímu signálu, nyní s odstraněnými komponentami s artefakty.
 - b. Opakujeme krok a., začátek posuneme o třetinu délky segmentu.
 - c. Opakujeme krok a., začátek posuneme o dvě třetiny délky segmentu.
3. Kompozice dílčích rekonstrukcí do výsledné
 - a. Pro každý z kanálů a každou z částí
 - I. Podle výše uvedeného postupu určíme, které signály se mají při skládání použít.
 - II. Výsledný signál v dané části je průměrem vybraných dílčích výsledků.

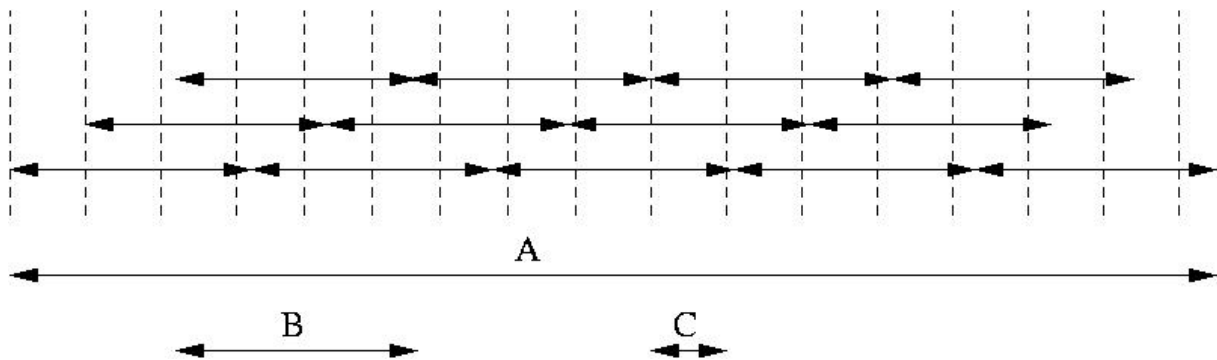


Diagram 1: Schema dělení dat (A) na segmenty (B), kde se provádí ICA, a dělení na části (C), kde se skládají tři dílčí rekonstrukce v jednu.

Aplikace algoritmu při odstraňování artefaktů z libovolně z dlouhého záznamu

Před samotnou aplikací algoritmu je nutné zvolit hodnoty volných parametrů.

K experimentálnímu určení mezní hodnoty pro anomalitu nám může posloužit obr. 3. Hodnoty anomaly v jednotlivých komponentách jsou v tab. 1.

komponenta	Anomalita
1	14.7318
2	24.0713
3	27.9069
4	12.7391
5	8.5842
6	76.2477
7	19.5738
8	12.6141

Tab. 1 – anomalita pro komponenty z obr. 3

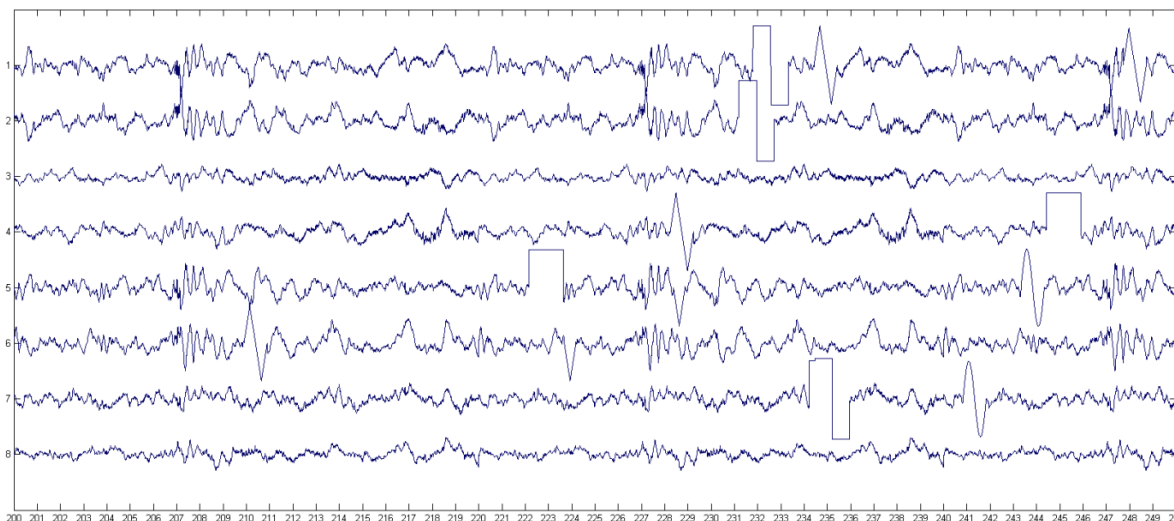
Výše uvedené hodnoty ukazují, že pro hodnoty nad 21 se v uvedeném příkladě za komponenty obsahující artefakty považují signály, které jsme do signálu sami přidali. Další závěry proto uvádíme vzhledem k takto zvolené mezní hodnotě.

Je dobré si všimnout kanálu číslo 7. Jeho průběh i ohodnocení je našim artefaktům nejpodobnější, přesto nemá s artefakty, které jsme do dat uměle vložili nic společného. Pravděpodobně se jedná o artefakt, který byl v datech přítomen již předtím. Těmito skutečnými artefakty se budeme zabývat v další části naší práce.

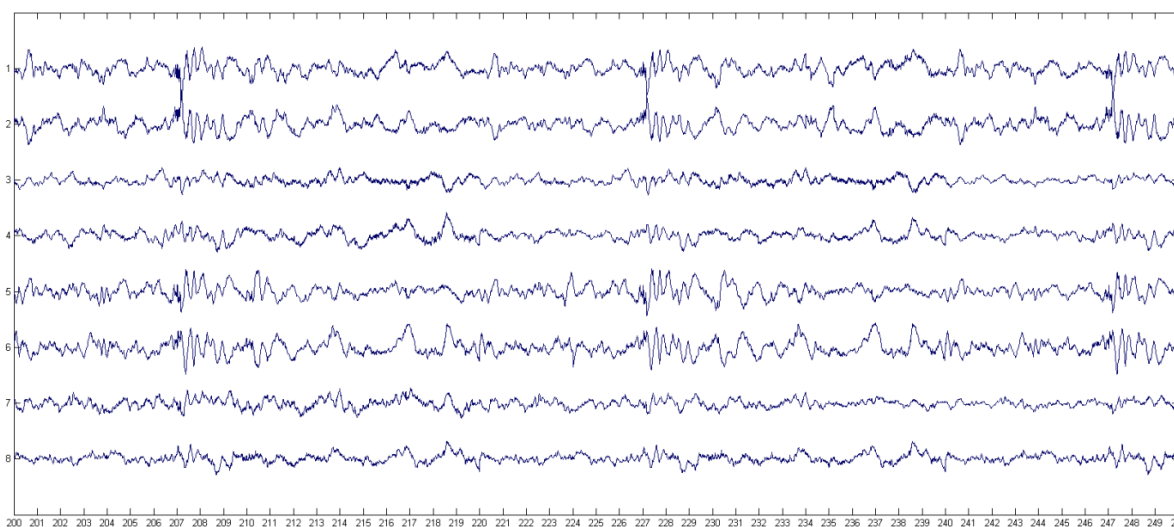
V tomto oddíle jsme hodnoty parametrů algoritmů nastavili jako

- a. Mezní hodnota pro anomalitu - 21
- b. Maximální počet odstraňovaných komponent - 6
- c. Metoda pro rozklad signálu na nezávislé komponenty - EFICA
- d. Délka segmentu pro analýzu signálu - 5120
- e. Délka částí pro rekonstrukci - 256

Při testování úspěšnosti navrženého algoritmu jsme do dat vložili náhodný počet výše uvedených artefaktů a následně se je snažili odstranit. Část dat s artefakty je na obr. 4, odpovídající část po aplikaci algoritmu je zobrazena na obr. 5.

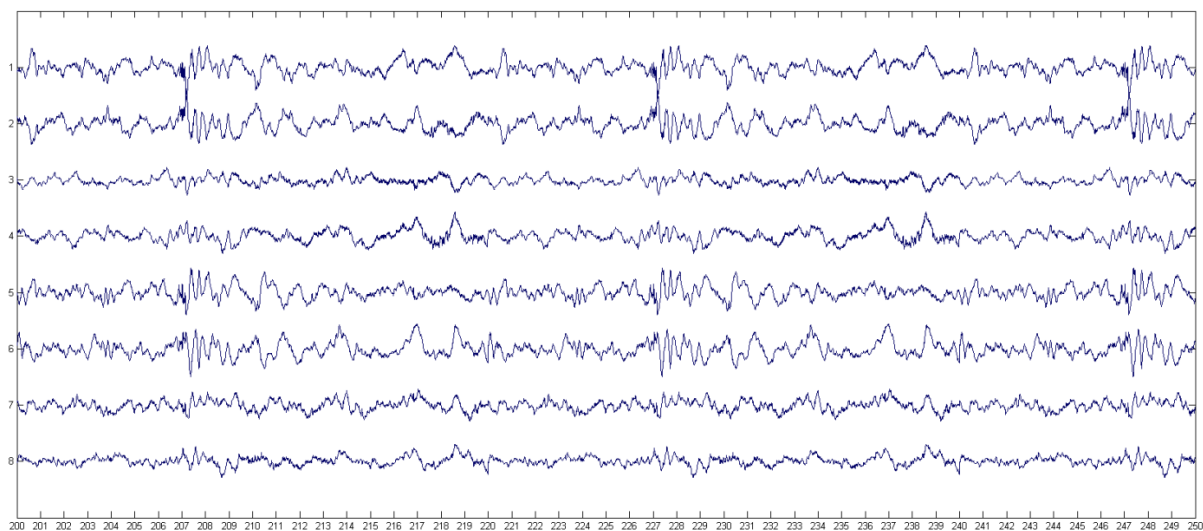


Obr. 4 – data s 13 artefakty (1. kanál – 232., 234. a 248 sekunda; 2. kanál – 231 sekunda; 4. kanál – 229. a 245. sekunda; 5. kanál – 222. sekunda, 228. a 244. sekunda; 6. kanál – 210. a 224. sekunda; 7. kanál - 223. a 241 sekunda)

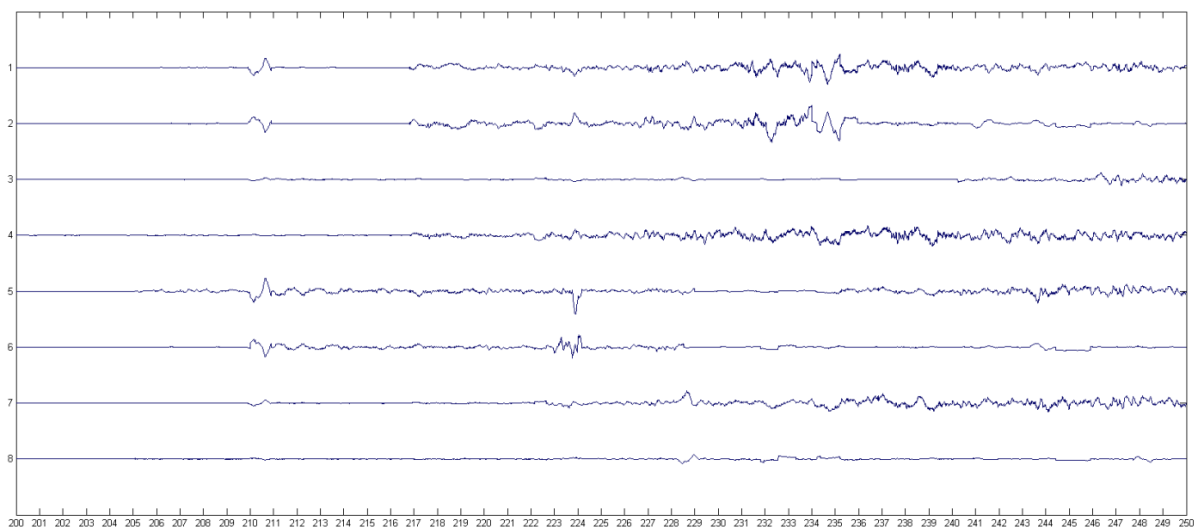


Obr. 5 – Vyčištěná data

Pro srovnání je na obr. 6 zobrazena stejná část signálu před přidáním artefaktů a na obr. 7 rozdíl tohoto signálu a signálu po vyčištění.

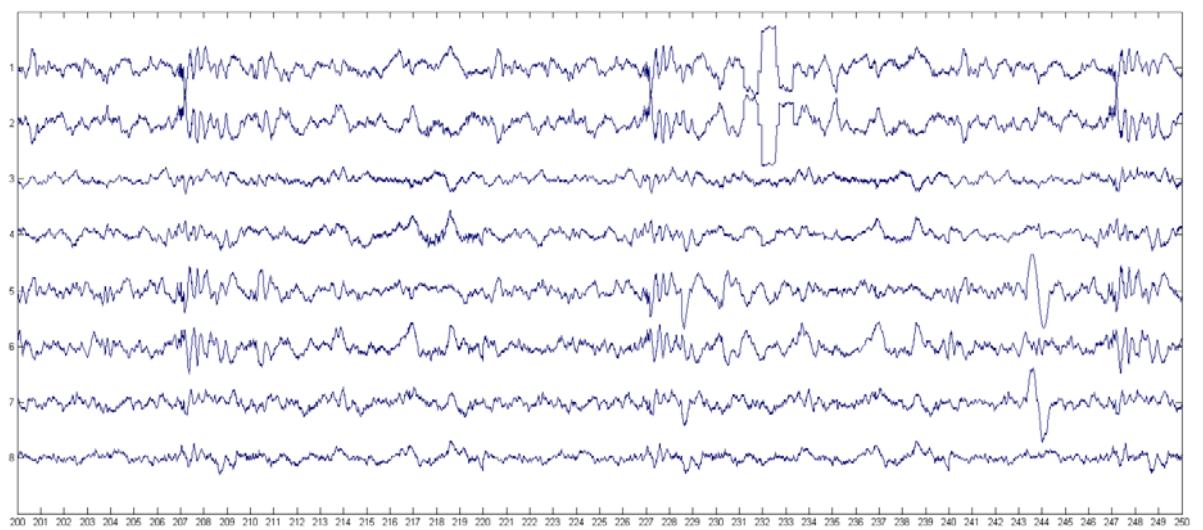


Obr. 6 – původní data

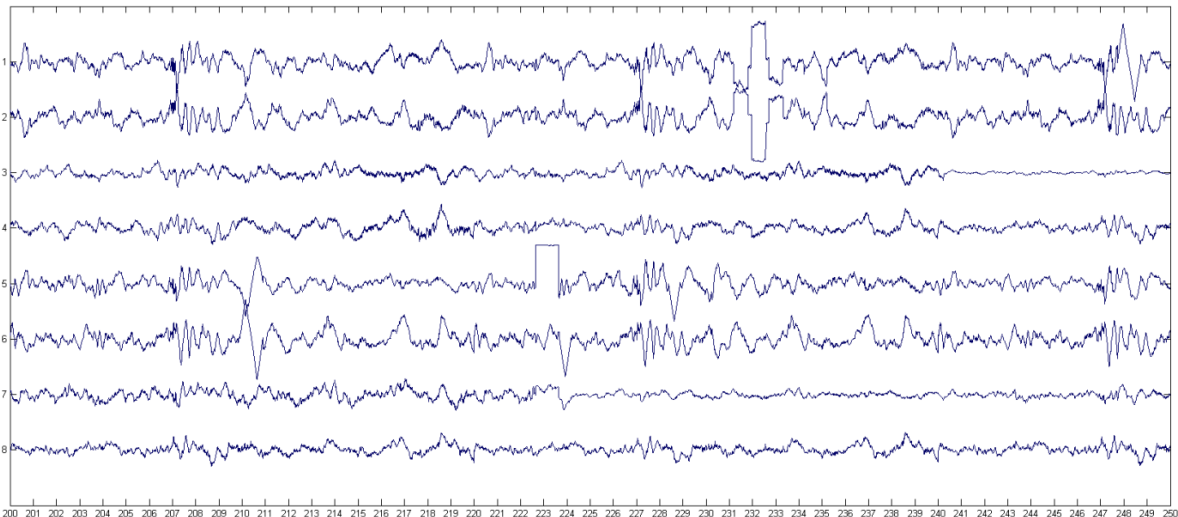


Obr. 7 – rozdíl původního signálu a signálu po vyčištění

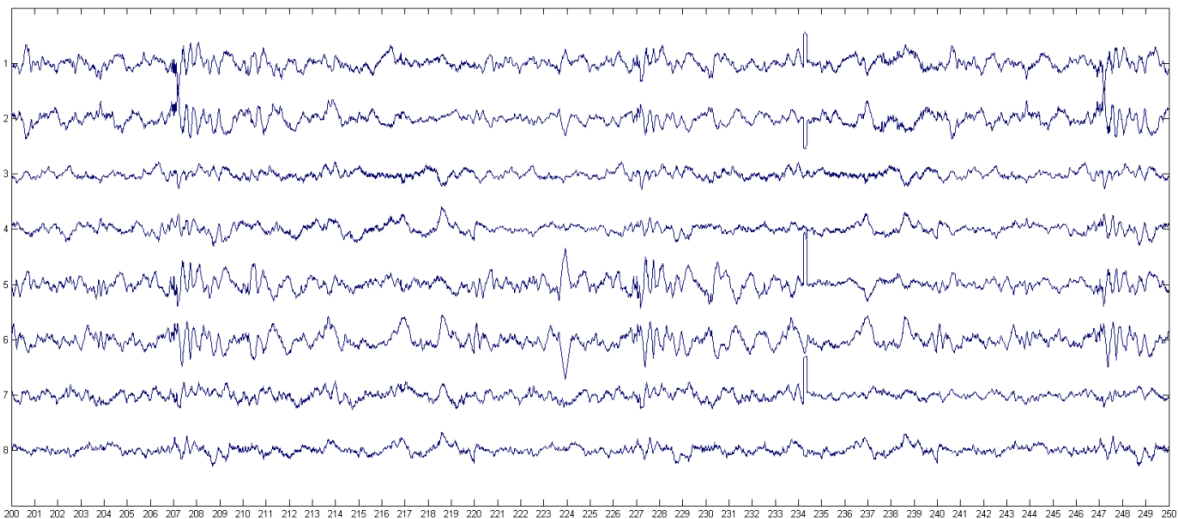
Pro posouzení vlivu komponování výsledného signálu z dílčí trojice připojujeme obrázky 8, 9 a 10, na kterých jsou zobrazeny dílčí signály.



Obr. 8 – data s částečně odstraněnými artefakty



Obr. 9 – data s částečně odstraněnými artefakty



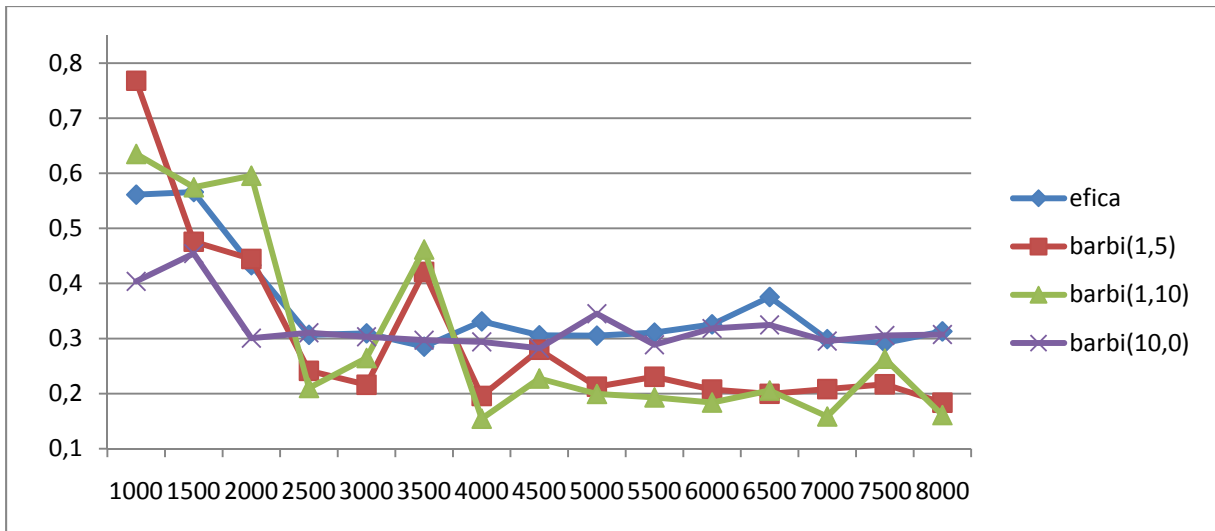
Obr. 10 – data s částečně odstraněnými artefakty

Srovnání metod pro rozklad signálu na nezávislé komponenty

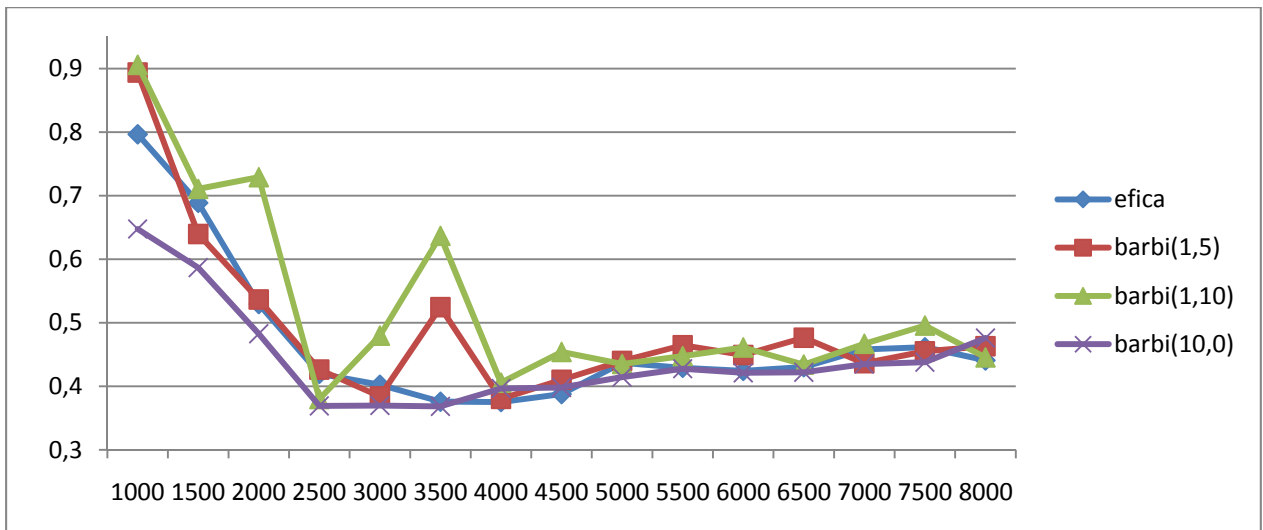
Pro určení optimální délky segmentu pro jednotlivé metody jsme vložili jeden, tři a pět artefaktů do dat tvořených jediným segmentem. Pokus jsme opakovali se dvěma různými záznamy, první byl snímán s frekvencí 256 Hz, druhý s frekvencí 128 Hz. Získané výsledky zobrazují grafy 1 až 6. Grafy zobrazují míry vyčištění dat na jejich délce. Za míru vyčištění je vzat poměr

$$\frac{\|x - z\|^2}{\|x - y\|^2}$$

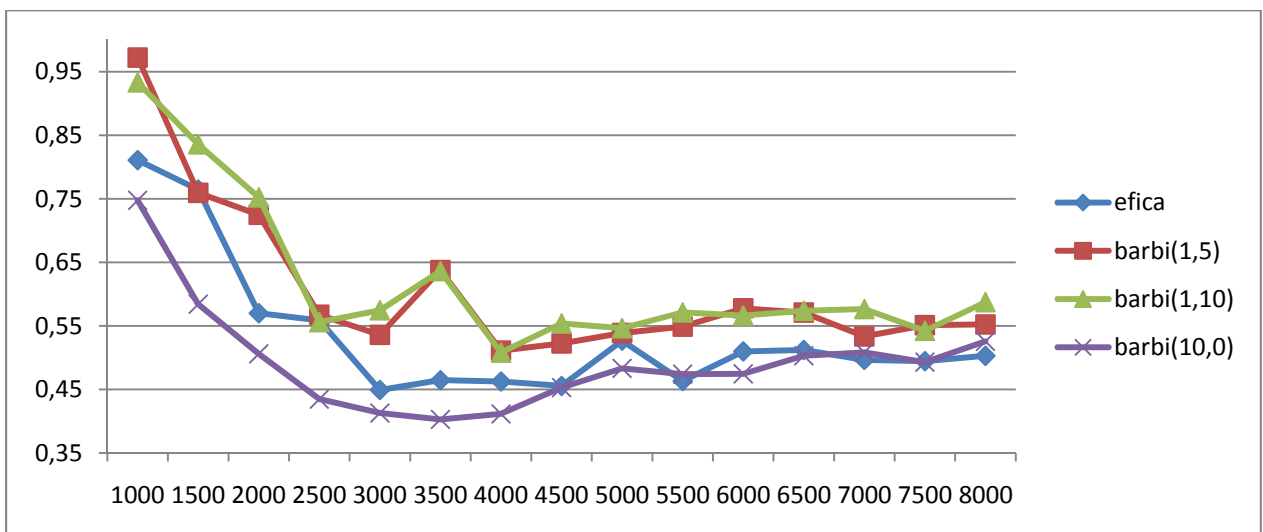
kde x jsou původní data, y data po přidání artefaktů a z data po vyčištění. Každý z pokusů byl opakován 100 krát.



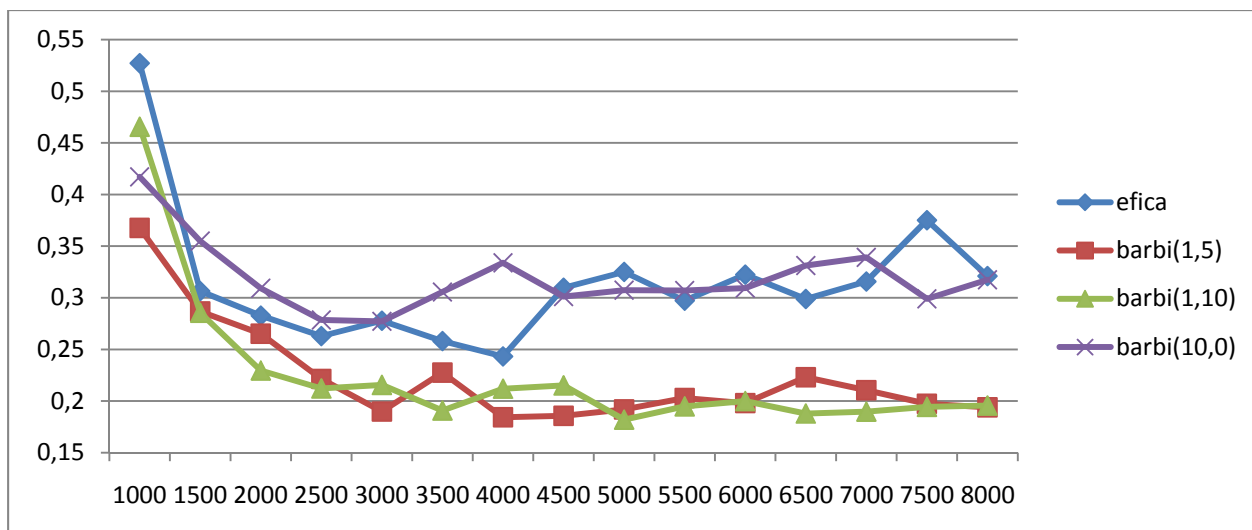
graf 1 - srovnání metod při vložení 1 artefaktu do dat snímaných s frekvencí 256 Hz



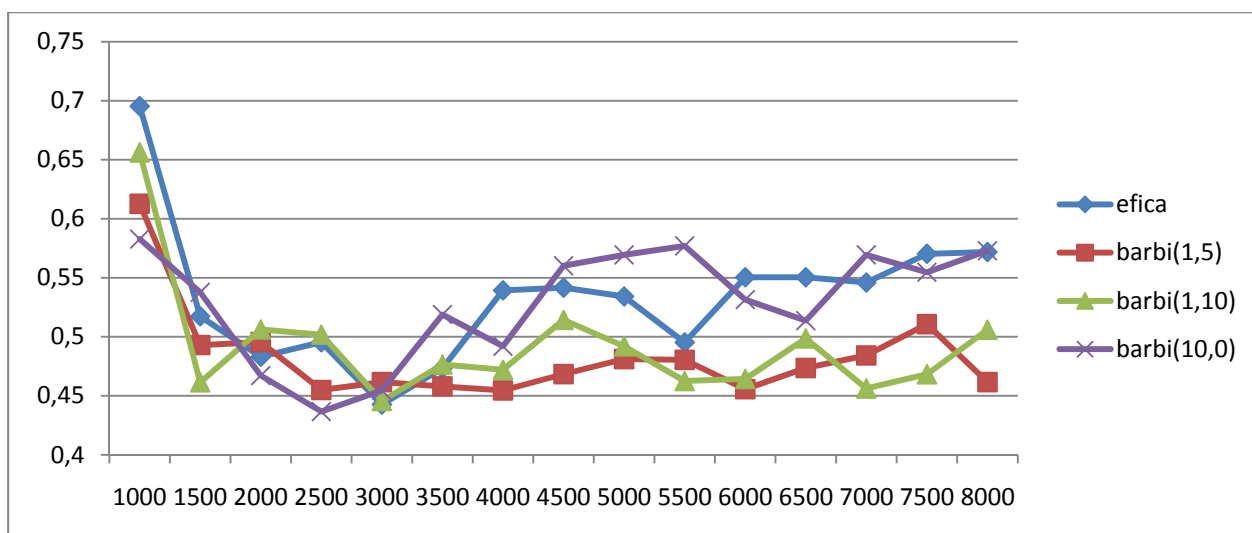
graf 2 - srovnání metod při vložení 3 artefaktů do dat s do dat snímaných s frekvencí 256 Hz



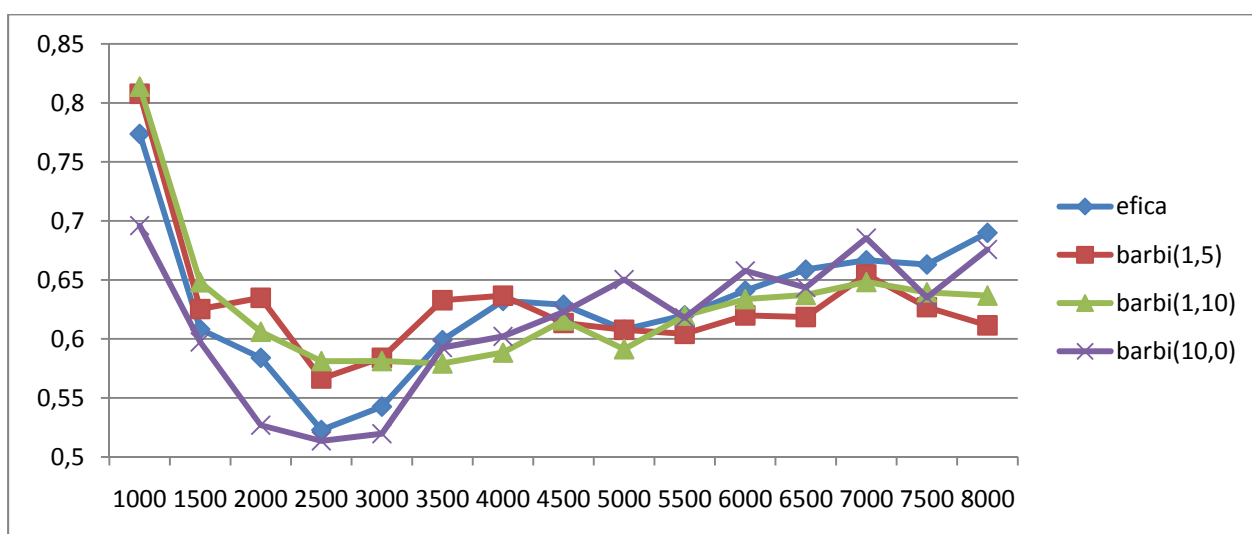
graf 3 - srovnání metod při vložení 5 artefaktů do dat snímaných s frekvencí 256 Hz



graf 4 - srovnání metod při vložení 5 artefaktů do dat snímaných s frekvencí 128 Hz



graf 5 - srovnání metod při vložení 5 artefaktů do dat snímaných s frekvencí 128 Hz



graf 6 - srovnání metod při vložení 5 artefaktů do dat snímaných s frekvencí 128 Hz

Optimální délku segmentu pro jednotlivé metody, zachycuje tab. 2. Při určování optimální délky jsme krom výsledků z předchozích grafů uvážili i fakt, že ve skutečných datech nevíme, jaké je

přesné rozložení artefaktů. Proto je rozumné volit spíše kratší segmenty, čímž omezíme možnost kumulace většího počtu artefaktů do jednoho segmentu.

Metoda	Optimální délka segmentu
EFICA	3000
BARBI(1,5)	4000
BARBI(1,10)	4000
BARBI(10,0)	2500

Tab. 2 - Optimální délka segmentu pro jednotlivé metody

Pro takto určené optimální délky segmentu jsme metody porovnávali na datech dlouhých 320 sekund, což při snímání s frekvencí 256Hz odpovídá 81920 datům. Do těchto dat jsme náhodně umístili 60 z výše uvedených artefaktů. Při použití stejných parametrů jsme pokus zopakovali se záznamem snímaným s frekvencí 128 Hz, který obsahoval 40960 dat. Každý z pokusů byl opakován 100 krát. Úspěšnost jednotlivých metod ukazuje tab. 3.

Metoda	Míra vyčištění dat snímaných s frekvencí 256 Hz	Míra vyčištění dat snímaných s frekvencí 128 Hz
EFICA	0.3753	0.4628
BARBI(1,5)	0.4033	0.5313
BARBI(1,10)	0.4132	0.5472
BARBI(10,0)	0.3557	0.4080

Tab. 3 – Srovnání metod

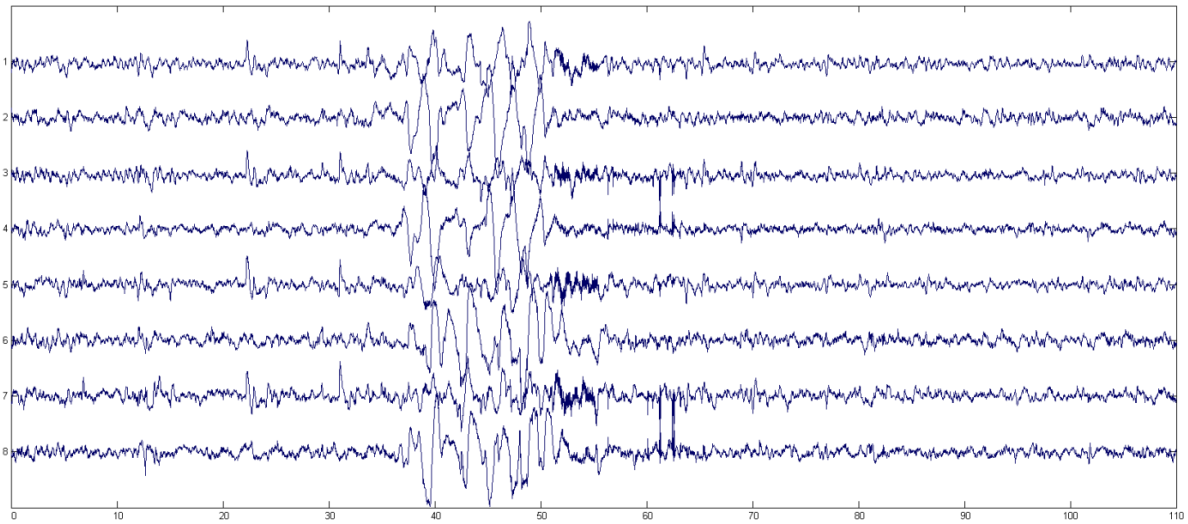
Aplikace algoritmu na data se skutečnými artefakty

Pro účely testování algoritmu na odstraňování skutečných artefaktů jsme použili část dat z EEG předčasně narozených dětí. Tyto data jsou zkoumány za účelem detekce spánkových fází. Všechny záznamy byly snímány s frekvencí 128 Hz.

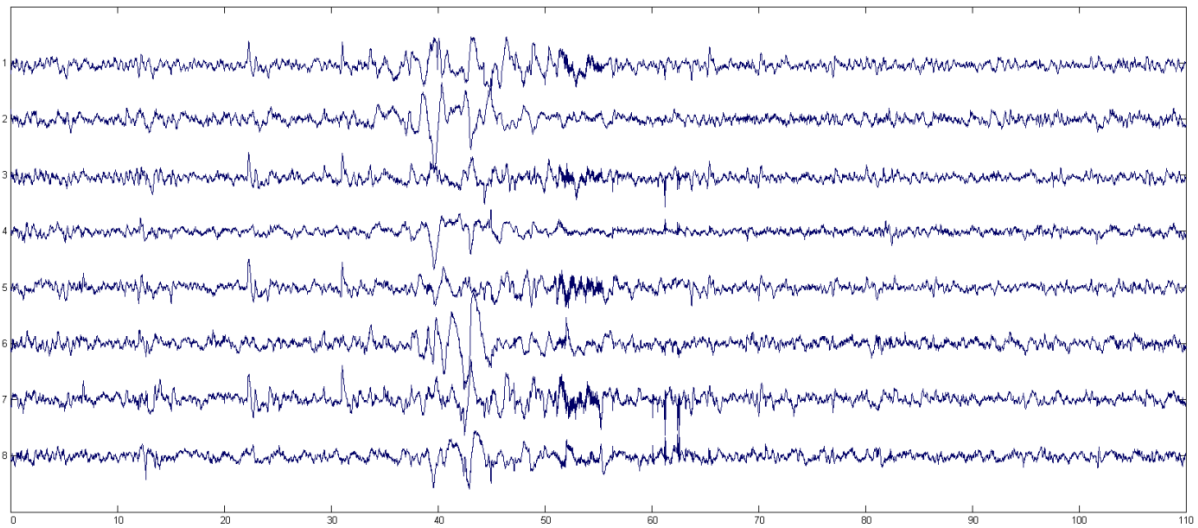
Při odstraňování artefaktů jsme, vzhledem k výsledkům z třetího oddílu, parametry algoritmu zvolili jako

- a. Mezní hodnota pro anomalitu – postupně 21,20,19,18 a 17
- b. Maximální počet odstraňovaných komponent - 6
- c. Metoda pro rozklad signálu na nezávislé komponenty - BARBI(10,0)
- d. Délka segmentu pro analýzu signálu - 2500
- e. Délka částí pro rekonstrukci – 256

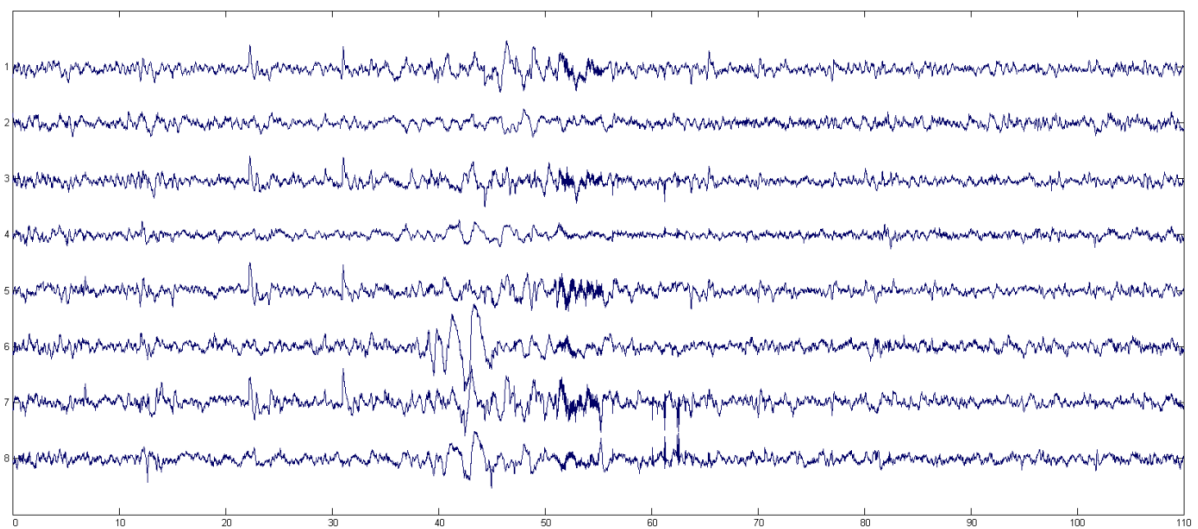
První zpracovávaná část dat s artefakty je na obr. 11. Výsledky po aplikaci algoritmu s jednotlivými volbami anomaly jsou na obr. 12 až 17.



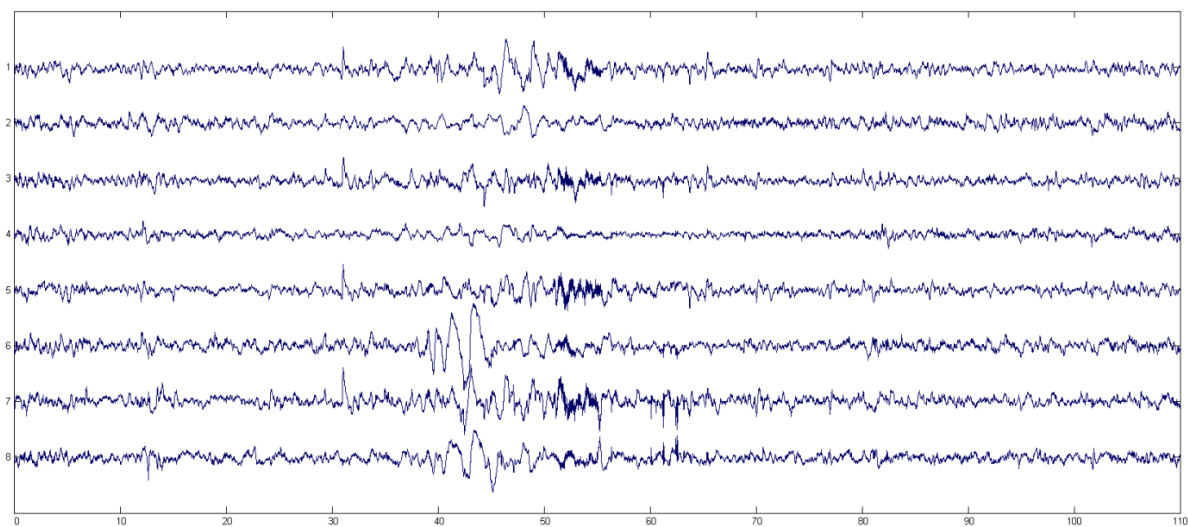
Obr. 11 – data s očními a pohybovými artefakty (operátorská poznámka „vrtí se“)



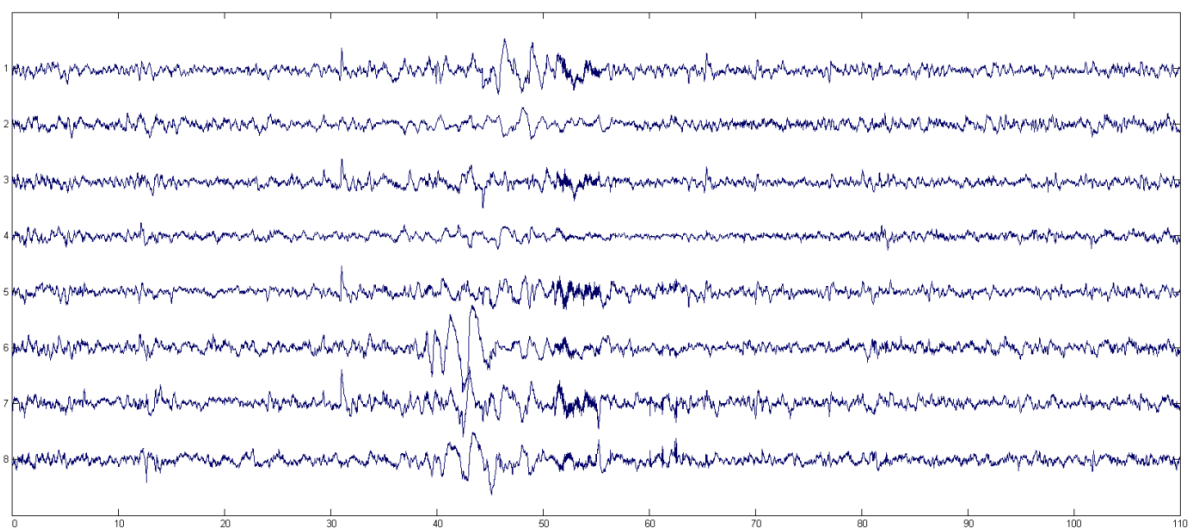
Obr. 12 – Mezní hodnota pro anomaliu=21



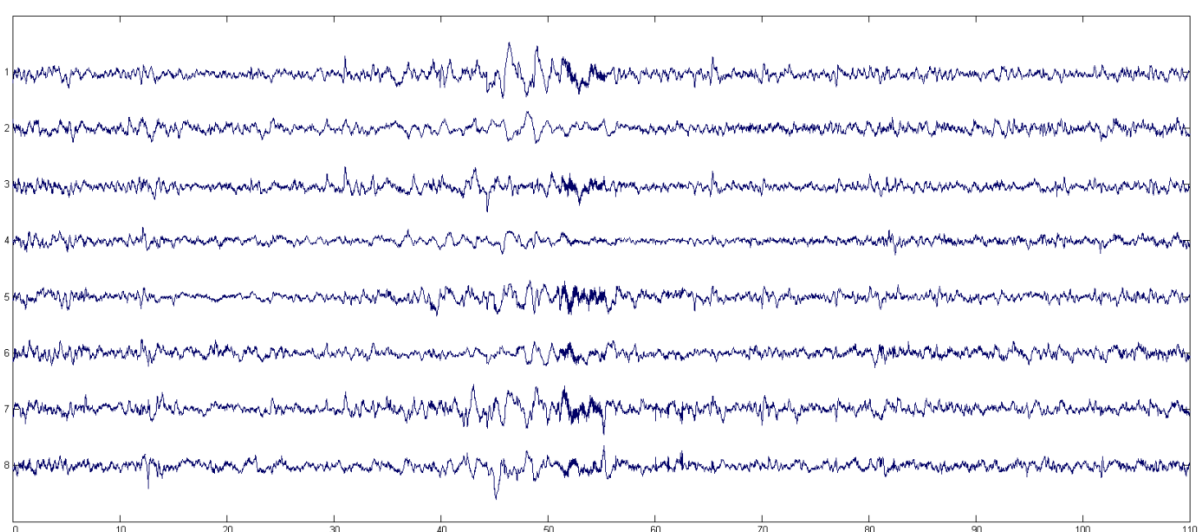
Obr. 13 – Mezní hodnota pro anomaliu=20



Obr. 14 – Mezní hodnota pro anomalitu=19

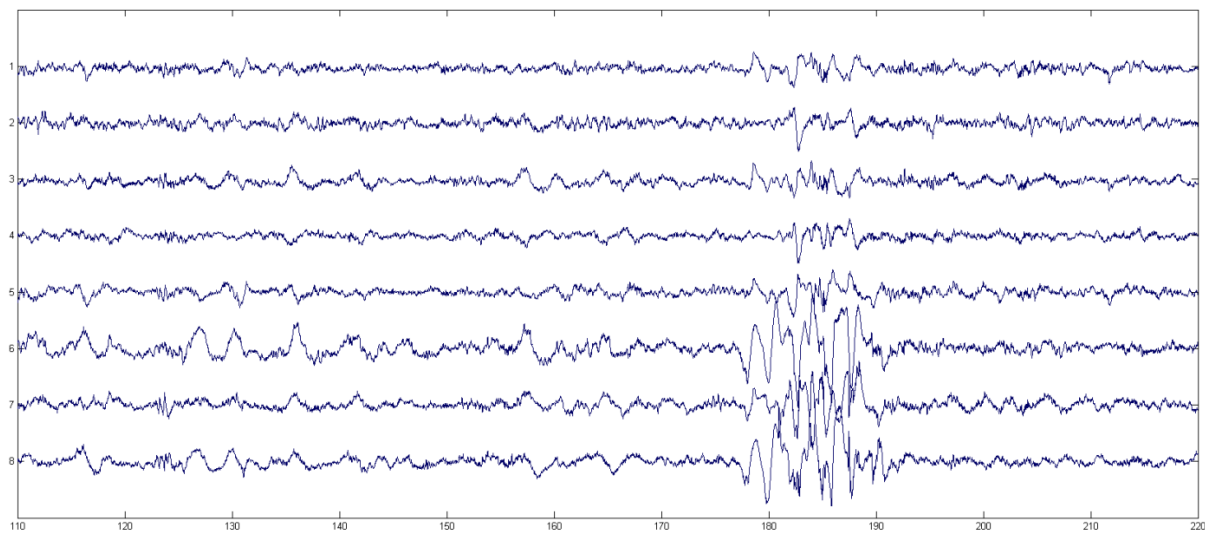


Obr. 15 – Mezní hodnota pro anomalitu=18

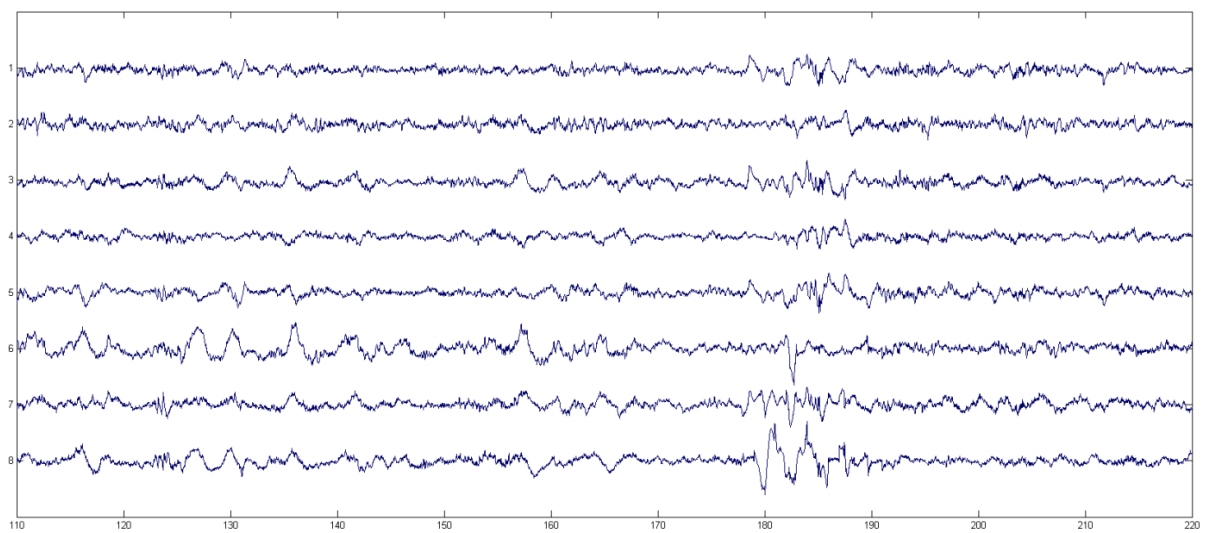


Obr. 16 – Mezní hodnota pro anomalitu=17

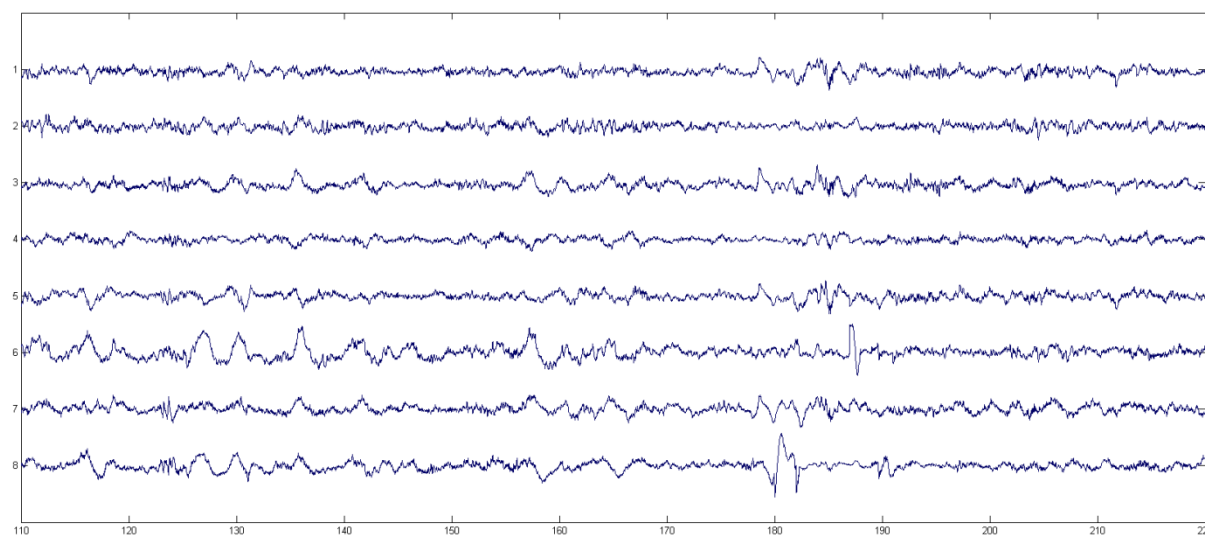
Druhá zpracovávaná část dat s artefakty je na obr. 17. Výsledky po aplikaci algoritmu s jednotlivými volbami mezní hodnoty pro označení artefaktu jsou na obr. 18 až 23.



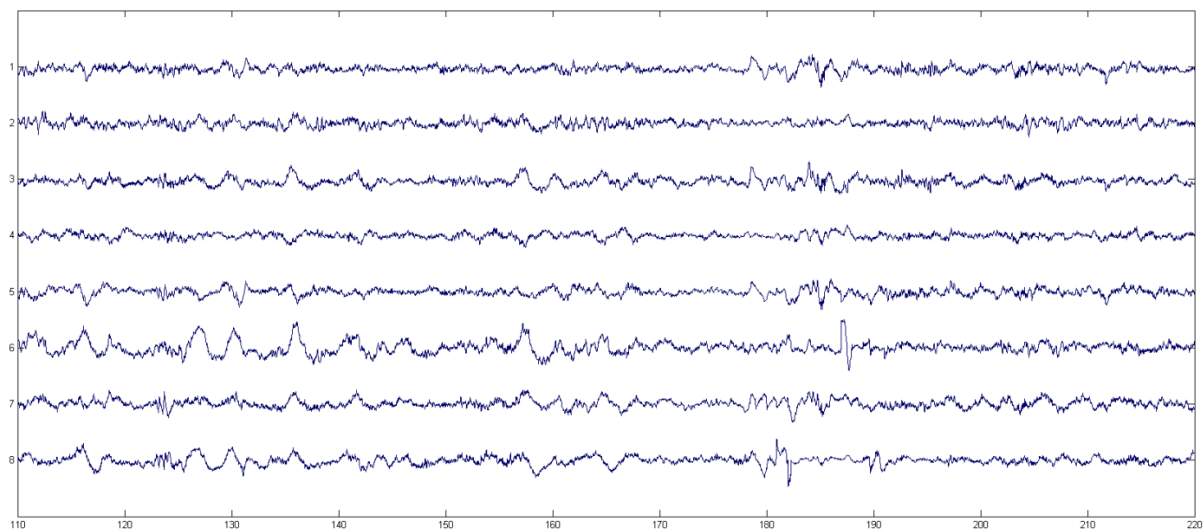
Obr. 17 – data s pohybovými artefakty



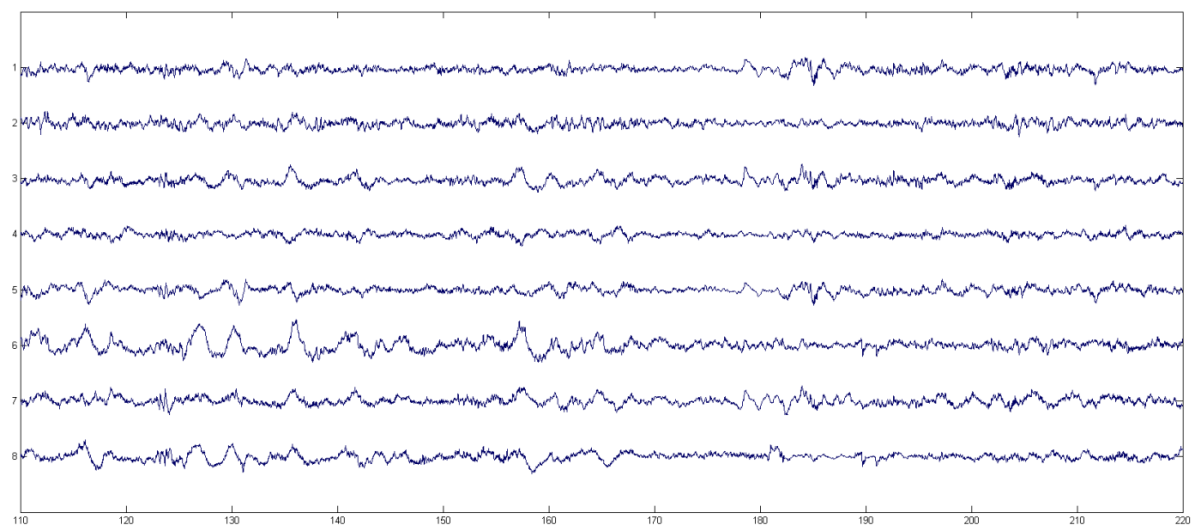
Obr. 18 – Mezní hodnota pro anomalitu=21



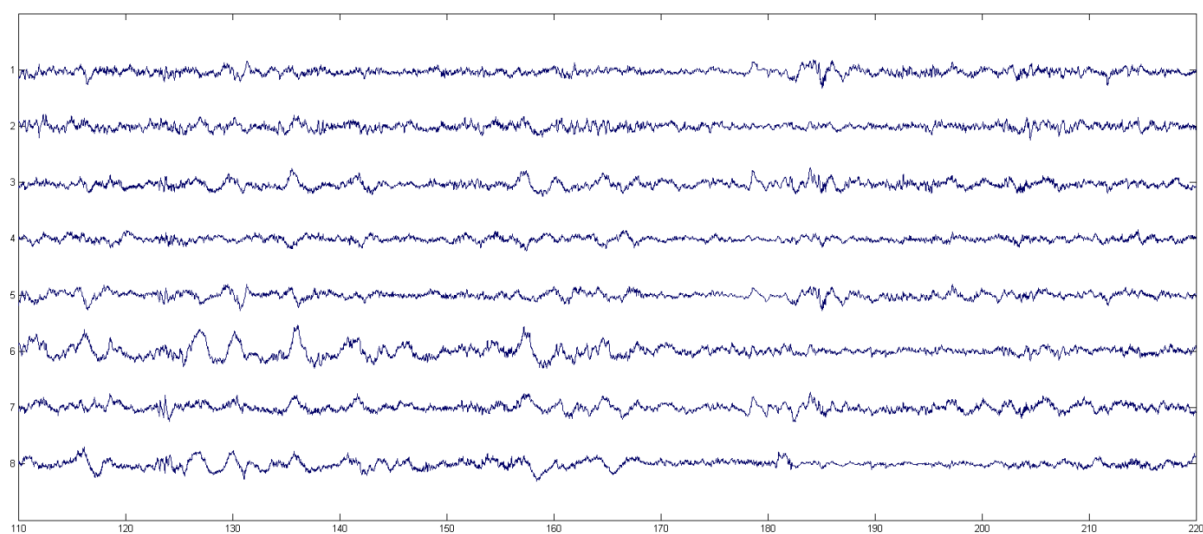
Obr. 19 – Mezní hodnota pro anomalitu=20



Obr. 20 – Mezní hodnota pro anomalitu=19

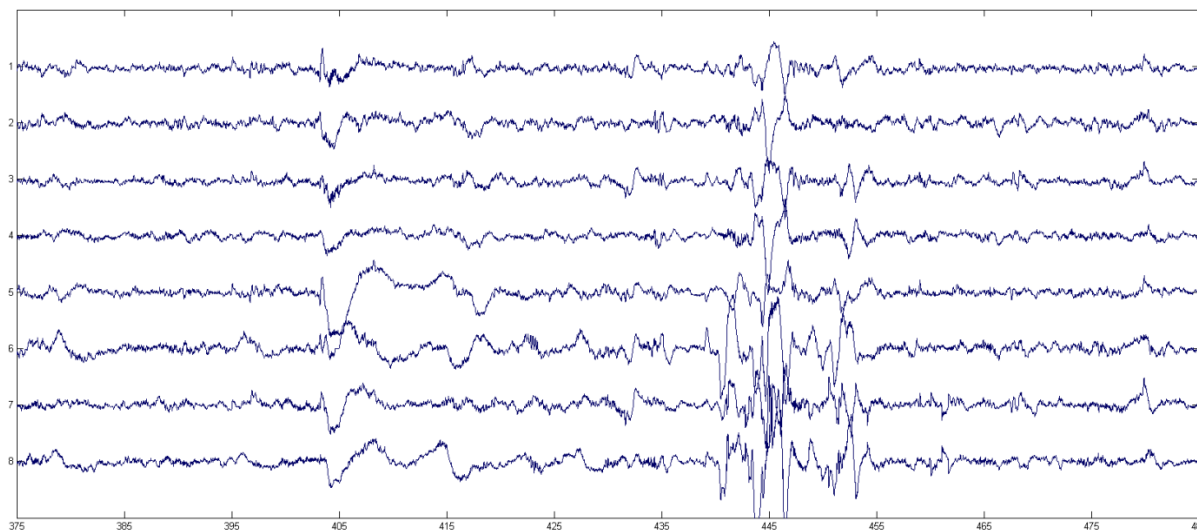


Obr. 21 – Mezní hodnota pro anomalitu=18

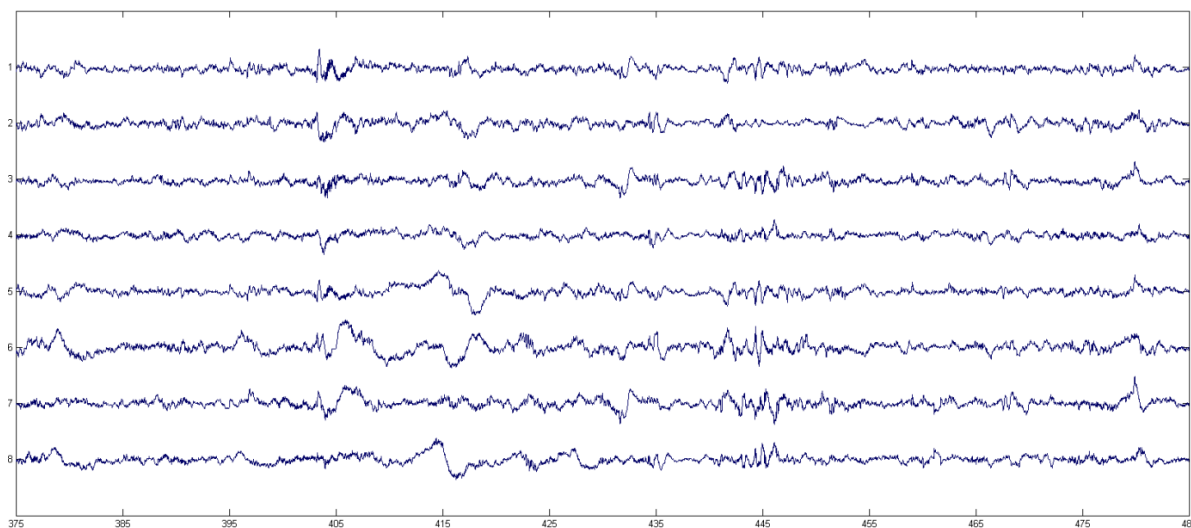


Obr. 22 – Mezní hodnota pro anomalitu=17

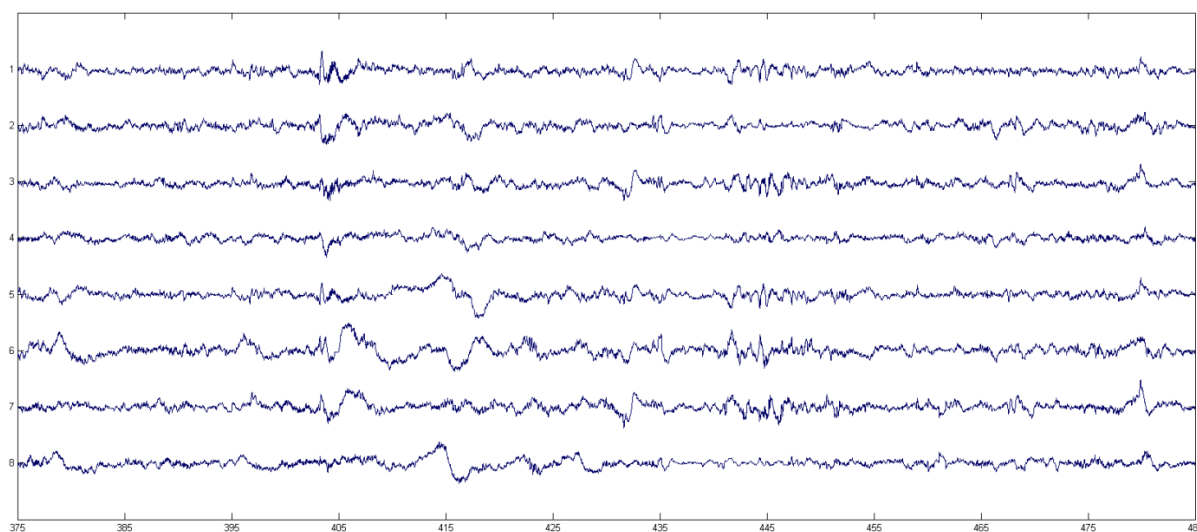
Třetí zpracovávaná část dat s artefakty je na obr. 23. Výsledky po aplikaci algoritmu s jednotlivými volbami mezní hodnoty pro označení artefaktu jsou na obr. 24 až 28.



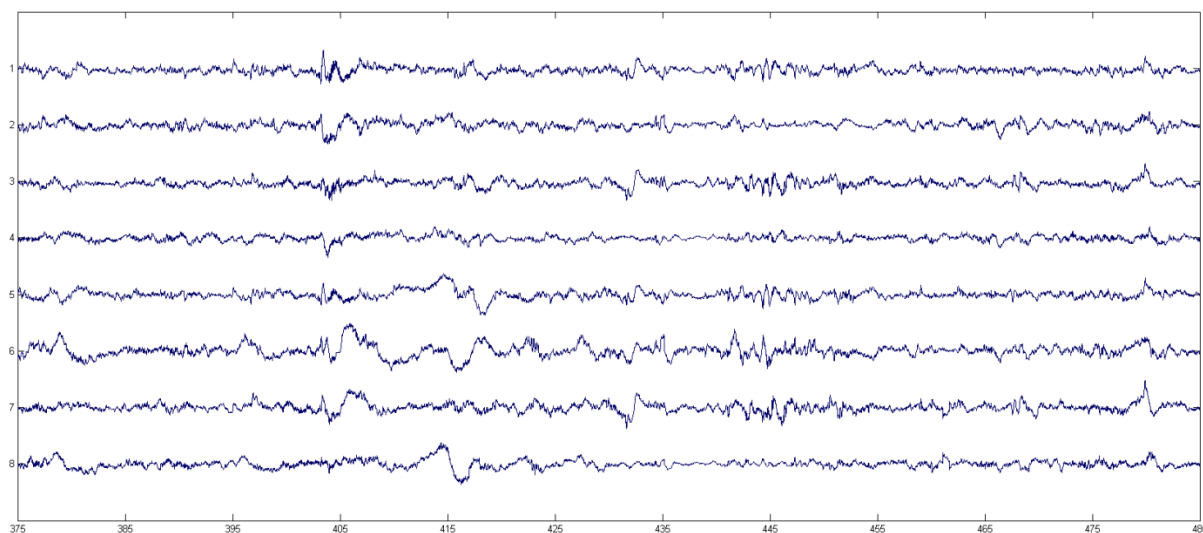
Obr. 23 – data s artefaktem typu odlepená elektroda a pohybovými artefakty



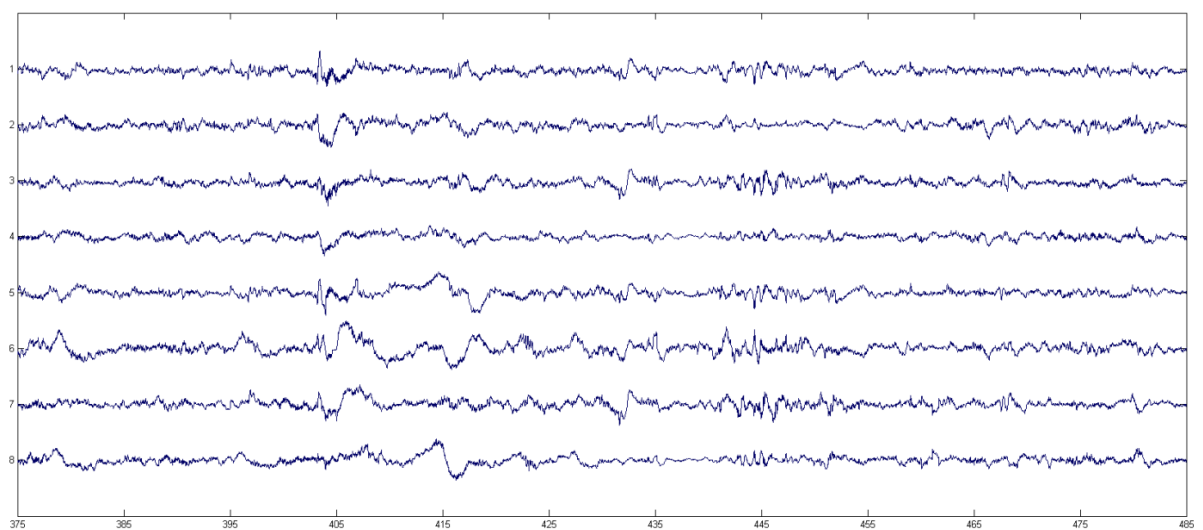
Obr. 24 – Mezní hodnota pro anomaliu=21



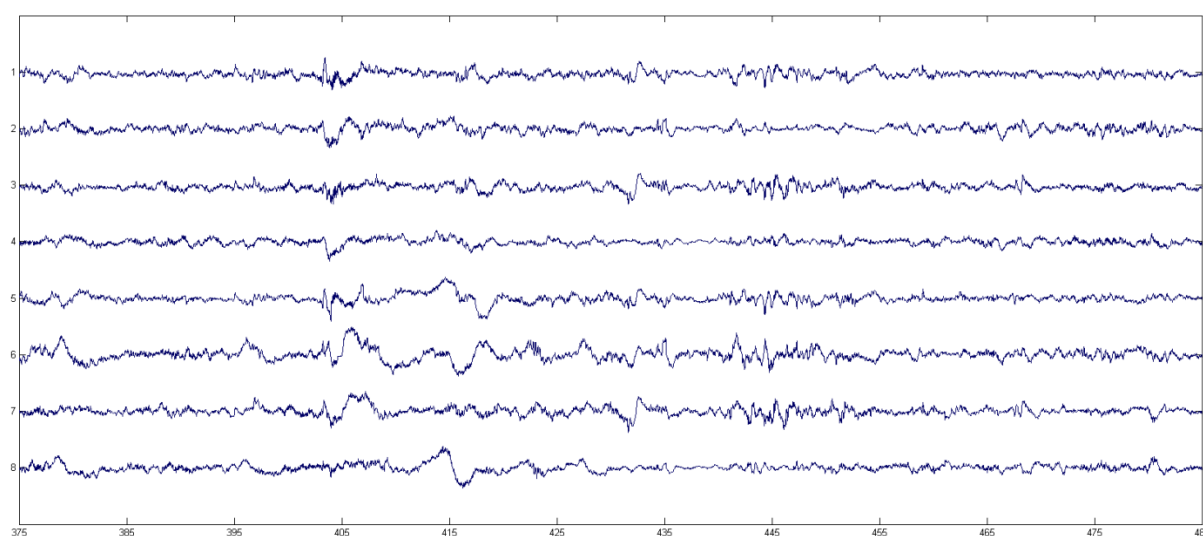
Obr. 25 – Mezní hodnota pro anomaliu=20



Obr. 26 – Mezní hodnota pro anomalitu=19



Obr. 27 – Mezní hodnota pro anomalitu=18



Obr. 28 – Mezní hodnota pro anomalitu=17

Při detekci spánkových stavů je pro správnou diagnózu rozhodující označení úseků, v nichž měřený subjekt prokazatelně vykazoval odlišnou mozkovou aktivitu. Určování těchto úseků

samozřejmě vzhledem k vysoké amplitudě artefakty znesnadňují. Po vyčištění signálu se, jak ukazují výše uvedené obrázky, artefakty buď zcela odstraní, nebo alespoň výrazně poklesne jejich amplituda. Navržený algoritmus pro odstraňování artefaktů by tedy mohl pomoci jako předzpracování dat při detekci spánkových stavů.

Závěr

V této práci byl představen algoritmus sloužící k odstraňování artefaktů v záznamech EEG. Jeho funkčnost jsme prokázali na reálných datech se simulovanými artefakty. K posouzení jsme rovněž předložili výsledky algoritmu na datech s reálnými artefakty.

Reference

- [1] Hyvarinen A., Karhunen J. and Oja E., A fast fixed point algorithm for independent component analysis, *Neural Computations*, 1997, 9, 1483-1492.
- [2] Belouchrani A., Abed-Meraim K., Cardoso J.F. and Moulines E., A blind source separation technique using second order statistics, *IEEE Tr. Signal Proc.* 1997, 45, 434-444.
- [3] Tang A.C., Sutherland M.T. and McKinney C.J., Validation of SOBI components from high-density EEG, *Neuroimage* 2004, 25, 539-553
- [4] Tichavský P., Nielsen J., Koldovský Z., "Odstraňování artefaktu v EEG datech I", výzkumná zpráva ÚTIA AV ČR č. 2140, říjen 2005.
- [5] Tichavský P., Nielsen J., Koldovský Z., "Odstraňování artefaktu v EEG datech II", výzkumná zpráva ÚTIA AV ČR č. 2176, prosinec 2006.
- [6] Koldovský Z., Tichavský P., Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the Cramer-Rao Lower Bound, *IEEE Tr. Neural Networks*, vol. 17, no. 5, pp. 1265-1277, September 2006.
- [7] Tichavský P., E. Doron, A. Yeredor and J. Nielsen, A Computationally Affordable Implementation of An Asymptotically Optimal BSS Algorithm for AR Sources, *EUSIPCO-2006*, Florence, Italy, September 4-8, 2006.
- [8] Tichavský P., Nielsen J., Krajča V., Identification of epileptic activity of brain in EEG recordings using four ICA techniques, *Proc. Biosignal 2006*, Brno, Czech Republic, June 28-30, 2006, pp.166-168.
- [9] Krajča V., S. Petránek, J. Mohylová, J. Nielsen, P. Tichavský, L. Lhotská, and M. Brunovský, Identification of neonatal EEG sleep stages by structural time profiles and independent component analysis, *Proc. WACBE World Congress on Bioengineering 2007*, Bangkok, Thailand. 9.-11.7. 2007.
- [10] Tichavský P. and A. Yeredor, Fast Approximate Joint Diagonalization Incorporating Weight Matrices, *IEEE Tr. on Signal Processing*, vol. 57, no.3, pp. 878-891, March 2009.
- [11] Tichavský P., A. Yeredor, and Z. Koldovský, A Fast Asymptotically Efficient Algorithm for Blind Separation of a Linear Mixture of Block-Wise Stationary Autoregressive Processes, *ICASSP 2009*, Taipei, Taiwan, April 21-27, 2009, pp. 3133-3136
- [12] Tichavský P., Matlab codes for EFICA, WASOBI, BARBI, <http://si.utia.cas.cz/Tichavsky.html>
- [13] Matlab toolbox EEGLAB, <http://sccn.ucsd.edu/eeglab/index.html>

Příloha 1: Matlabové funkce, které realizují daný návrh

```

function eegout=cleareeg(eegin,meza,maxo,met,seq,len)

%Input data:
%eegin ... input EEG data
%meza ... cut off point for anomaly of independent component to be deleted
%maxo ... maximum number of deleted components
%met ... method for evaluating demixing matrix
%     possible options - 'e' ... EFICA
%     'b5' ...BARBI(1,5)
%     'b10' ...BARBI(1,10)
%     'b10,0' ...BARBI(10,0)
%seq ... lenght of segments
%len ... lenght of part of which is composed clean EEG from partially cleaned ones

%Output data:
%eegout...Output EEG data

%Default values of parameters
if nargin < 6
    len=256;
    if nargin < 5
        seq=2500;
        if nargin < 4
            met='b10,0';
            if nargin < 3
                maxo=6;
                if nargin < 2
                    meza=17;
                end
            end
        end
    end
end
end
end

delka=size(eegin,2);

if delka/seq>1 %if lenght of eeg is enough for dividing into segments
    presah=floor(seq/3);

    A=simplerepclean(eegin,meza,maxo,met,seq); %partially cleaned eeg1

    B(:,1:presah)=A(:,1:presah);
    B(:,presah:delka)=simplerepclean(eegin(:,presah:delka),meza,maxo,met,seq); %partially cleaned eeg2

    C(:,1:(seq-presah))=A(:,1:(seq-presah));
    C(:,(seq-presah):delka)=simplerepclean(eegin(:,(seq-presah):delka),meza,maxo,met,seq); %partially cleaned eeg3

    %calculation of interval between averange signal and zero signal
    mez=0;
    D=zeros(1,len+1);
    for i=1:10
        S=floor((delka-len)*rand);
        mez=mez+norma(A(randi([1,size(eegin,1)]),S:S+len),D);
    end;
    mez=mez/5;

    %composition of clean eeg from partialy cleaned eeg
    poc=floor(delka/len); %number of parts in each channel
    zac=1;
    for i=1:poc
        if i==poc
            kon=delka;
        else
            kon=zac + len;
        end

        for j=1:size(eegin,1)
            a=A(j,zac:kon);
            b=B(j,zac:kon);
            c=C(j,zac:kon);

            %calculation of intervals between part of signals
            info(1)=norma(a,b);

```



```

end

pomdat=eegout(:,zac:kon); %part of eeg in given segment

%choice of method for evaluating demixing matrix
switch met
case('e')
    W=efica(pomdat);
case('b5')
    W=barbi(pomdat,1,5);
case('b10')
    W=barbi(pomdat,1,10);
case('b10,0')
    W=barbi(pomdat,10,0);
end

pomdat=W*pomdat; %independent components
cena=ocena(pomdat); %values of criterion for each components

%delete components with values of criterion above cut off point
if (max(cena)>meza)
    for j=1:maxo
        [cenamax,imax]=max(cena);
        if (cenamax>meza)
            pomdat(imax,:)=0;
            cena(imax)=0;
        end
    end
    eegout(:,zac:kon)=inv(W)*pomdat; %part of partialy cleaned eeg in given segment
end;
zac=kon;
end
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function norm=norma(x,y)

%Input data:
%x ... signal1
%y ... signal2

%Output data:
%info ... interval between x and y

    pom=(x-y).^2;
    norm=sqrt(sum(pom(:)));
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function info=ocena(y)

%Input data:
%eegin ... independent component

%Output data:
%info ... values of criterion

    m=max(y');
    y=((y./(m*ones(1,size(y,2))))*1.5).^10;
    info=log((mean(y,2))./(median((y),2)));
end

```

Poděkování:

Tato práce vznikla za podpory Grantové agentury ČR, project 102/09/1278, a Ministerstva školství, mládeže a tělovýchovy, project 1M0572.