



VILNIAUS UNIVERSITETAS
MATEMATIKOS IR INFORMATIKOS FAKULTETAS
KOMPIUTERIJOS KATEDRA

Magistro baigiamasis darbas

Medicininės kilmės signalų analizė

Atliko:

Vytenė Masytė

parašas

Vadovas:

doc. Tadas Meškauskas

Vilnius
2017

Turinys

Santrauka	4
Summary	5
Iyadas	6
1. EEG svarba ir taikymas	8
1.1. EEG diagnostika	8
1.2. Smegenų veikla ir epileptofominių pikų požymiai	9
2. EEG pikų aptikimas	10
2.1. Susijusių darbų apžvalga	10
2.2. EEG pikų aptikimo algoritmas, remiantis morfologiniais filtrais	10
2.2.1. Morfologinių filtrų teorija	10
2.2.2. Struktūrinio elemento parinkimas	12
2.2.3. EEG signalo tiriamos atkarpos ilgio parinkimas	13
2.2.4. Struktūrinio elemento ilgio parinkimas	14
2.2.5. Pikų detekcijos riba	15
2.3. EEG pikų aptikimo algoritmas, remiantis interpoliavimu gabalais tiesine funkcija	15
2.3.1. Gabalais tiesinė funkcija	15
2.3.2. Teorinis algoritmo aprašymas	16
2.3.3. Gabalais tiesinės funkcijos skaičiavimas	17
2.3.4. Mažiausių kvadratų metodo naudojimas tiesių radimui	17
2.3.5. Tiesių koeficientų reikšmė	18
3. Pagalbiniai įrankiai algoritmų įgyvendinimui	20
3.1. Testinių duomenų bazė	20
3.1.1. Programoje naudotos bibliotekos	20
3.2. Naudoti bei sukurti įrankiai analizei	21
3.2.1. Triukšmo šalinimas signaluose	22
4. Morfologinių filtrų algoritmo įgyvendinimas	24
4.1. Struktūrinio elemento ilgio parinkimas	24
4.2. Struktūrinio elemento aukščio parinkimas	25
4.3. Struktūrinių parabolinių skirtumas	26
4.4. Pikų detekcijos ribos skaičiavimas	27
4.4.1. Encefalogramų tyrimui parinkimas	28
4.4.2. Detekcijos ribos koeficiento tikslinimas	28
5. Interpoliavimo gabalais tiesine funkcija algoritmo įgyvendinimas	30
5.1. Intervalo pasirinkimas interpoliavimui	30
5.2. Žingsnio pasirinkimas interpoliavimui	31
5.3. Epileptofominių pikų detekcija	32
5.3.1. Tiesių koeficientų svarba	32
5.3.2. Tiesių koeficientų skirtumo parinkimas	33
5.3.3. Encefalogramų tyrimui parinkimas	34

5.3.4. Koeficiento m tikslinimas	34
6. Algoritmų palyginimas ir rezultatai	36
6.1. Algoritmų patikimumo vertinimas	36
Išvados ir rekomendacijos	38
Ateities tyrimų planas	40
Literatūros šaltiniai	41

Santrauka

Šiame darbe yra analizuojamos elektroencefalogramos (EEG), kurios yra pagrindinis įrankis epilepsijos ir kitų centrinės nervų sistemos ligų diagnozavime. Epilepsijos pikų nustatymui buvo įgyvendinti du algoritmai, bei atlikta jų analizė. Pirmasis algoritmas remiasi morfologinių filtrų teorija. Antrasis interpoliavimu gabalais tiesine funkcija. Pagrindinis šio darbo tikslas yra algoritmuose naudojamų koeficientų tikslinimas, bei šių algoritmų tarpusavio palyginimas. Morfologinių filtrų algortime naudojami koeficientai yra skirti struktūrinio elemento konstravimui, kuris apibūdina foninę smegenų veiklą ir leidžia morfologiniams filtrams išryškinti pikus, nufiltruojant foninius smegenų ritmus. Tikslesniam ir greitesniam koeficientų radimui, buvo sukurta grafinė sąsaja, kuri leidžia vizualiai aptikti optimaliausią variantą. Interpoliavime gabalais tiesine funkcija, tiesių koeficientai apibūdina pikų kampą, detalus jų nustatymas leidžia tiksliau detektuoti pikus.

Summary

Analysis of medical signals

The aim of this work is to analyse electroencephalograms (EEG) which is the main tool in diagnosis and treatment of epilepsy. Two algorithms which could detect epileptic spikes are presented in this work. First is based on morphological filters theory and the second one on piecewise linear curve fitting. The main result of this work is to compare two mentioned algorithms and find the correct values for coefficients used in their execution. For morphological filters, structuring element and its calculation was described. For an efficient results achievement a special tool was created, which allowed to see graphical change of the signal if any of value was increased or decreased. For piecewise linear fitting algorithm, the slope constants of the linear equations were calculated which helped to detect epileptic spikes. All implementations were tested on special datasets in which the count of spikes were already known.

Iyadas

Medicininės kilmės signalai, tai signalai, aprašantys žmogaus organizmo veiklą ir įvertinantys sveikatos būklę. Šiame darbe pagrindinis dėmesys yra skiriamas smegenų veiklos tyrimui. Elektroencefalografija¹ - tai smegenų biosrovės matavimo metodas, elektrinių virpesių fiksavimas. Gautųjų elektrinių smegenų aktyvumo signalų užrašymas vadinamas elektroencefalograma (EEG). Iš EEG galima spręsti apie smegenų aktyvumą bei sąmonės būklę, todėl šis metodas plačiai naudojamas centrinės nervų sistemos ligų, epilepsijos diagnostavime.

Medicininės kilmės signalų analizė, kuomet gydytojas rankiniu būdu stebi užrašytą signalą, vis dar dažnai yra taikoma medicinoje. Epilepsijos ligos atveju, signale yra ieškoma tam tikrų pikų, vadinamų epileptoforminiais pikais. Gydytojui vizualiai aptikti šiuos pikus - reikalaujantis didelių įgūdžių ir daugybės laiko darbas. Todėl nenuostabu, jog pilnai kompiuterizuoti šių pikų paiešką, yra daugiau nei 40 metų besitęsiantis tikslas [6]. Šiame darbe yra aprašomi algoritmai, kurių pagalba, galima automatizuoti epileptoforminių pikų aptikimą ir taip pagelbėti gydytojams, epilepsijos diagnostikoje.

Magistro baigiamojo darbo tikslai:

- Susipažinti su EEG pikų aptikimo teorija ir jau įgyvendintais moksliniais darbais šioje srityje.
- Įgyvendinti algoritmą, paremtą morfologinių filtrų teorija.
- Įgyvendinti algoritmą, paremtą interpoliavimu gabalais tiesine funkcija.
- Nustatyti skaičiavimo koeficientų vertes, su kuriomis įgyvendinti algoritmai veikia tiksliausiai EEG pikų detekcijai, turimoje signalų duomenų bazėje.
- Palyginti įgyvendintuosius algoritmus tarpusavyje, vertinant jų pikų detektavimo rezultatus.

Pirmajame darbo skyriuje aprašoma EEG tyrimo svarba ir taikymas medicinoje, pateikiamos epileptoforminio piko savybės.

Antrasis skyrius skirtas morfologinių filtrų bei algoritmo, paremto interpoliavimu gabalais tiesine funkcija, teorijai bei aprašymams. Taip pat nagrinėjami EEG pikų aptikimo srityje jau atlikti kitų autorių darbai.

Trečiajame skyriuje pateikiami pagalbinių algoritmų ir analizei sukurtų įrankių aprašymas bei veikimas. Aprašoma testinių duomenų bazė bei patys duomenys.

Ketvirtasis ir penktasis skyriai - tai praktinė signalų analizė: įgyvendintų algoritmų ir jiems nagrinėti sukurtais įrankiais gautų rezultatų aptarimas bei algoritmuose naudojamų koeficientų tikslinimas. Šeštajame skyriuje pateikiamas algoritmų tarpusavio palyginimas, formuluojamos išvados apie algoritmų veikimo kokybę.

Šis magistro baigiamasis darbas yra tęstinis, jame yra medžiagos iš praėjusio semestro mokslo tiriamojo darbo. Mokslo tiriamajame darbe, buvo nagrinėjamas morfologinių filtrų algoritmo veikimas, todėl epileptoforminių pikų aprašymas bei savybės, mokslinių darbų, susijusia su šia tema nagrinėjimas bei morfologinių filtrų teorija buvo aprašyti praėjusio semestro metu. Tačiau morfologinių filtrų koeficientų tikslinimas, epileptoforminių pikų detekcijos ribos nustatymas bei

¹gr. encephalos – smegenys

pikus detektuojančio algoritmo įgyvendinimas, remiantis morfologiniais filtrais, buvo pilnai atlikti šiame magistro baigiamajame darbe.

Interpoliavimo gabalais tiesine funkcija algoritmo aprašymas, koeficientų tikslinimas bei įgyvendinimas yra nauji darbai, atlikti šio semestro metu.

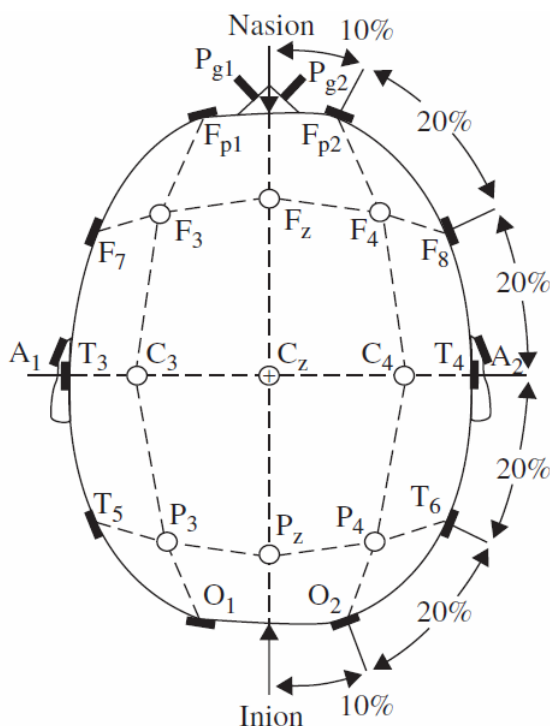
1. EEG svarba ir taikymas

Pasaulinės sveikatos organizacijos duomenimis, 50 milijonų pasaulio gyventojų serga įvairiomis epilepsijos formomis, todėl šios ligos diagnozavimas ir tyrimai turi didelę reikšmę medicinoje. Pirminis ir gana dažnas atliekamas tyrimas, įtariant epilepsiją, yra galvos smegenų encefalografija. Jos metu, yra fiksuojama smegenų biosrovė, užrašomas jų aktyvumas. Gydytojai encefalogramose ieško neįprastos smegenų veiklos, jų suaktyvėjimo, o epilepsijos atveju, vadinamųjų epileptoforminių pikų.

1.1. EEG diagnostika

Galvos smegenų EEG atliekama naudojantis paviršiaus elektrodais, juos tvirtinant prie galvos odos. Plačiausiai pasaulyje šiuo metu yra naudojama "10-20 sistema" elektrodų išdėstymui [12]. Pagal šią sistemą, kiekvienas elektrodas turi standartizuotą vietą, tai reiškia, jog jis yra tvirtinamas visada tuose pačiuose paciento galvos taškuose. "10" ir "20" apibūdina 10 % arba 20% atstumus tarp išdėstytų elektrodų. Paprastai EEG metu, tvirtinamas 21 elektrodas, kuris fiksuoja smegenų aktyvumą, savo matavimo kanale.

Ambulatoriniams ligoniams EEG trukmė yra apie 20 minučių, tačiau stacionariai gydomiems pacientams, dažnai atliekami detalesni ir ilgesni tyrimai, kurių signalų ilgis gali būti keliolika valandų. Buvo įrodyta, jog ilgesnės trukmės tyrimai, turi įtakos epilepsijos detekcijai [11], todėl vis dažniau prailgintos trukmės tyrimai gali būti atliekami ir ambulatorinėmis sąlygomis.



1 pav. Standartinės "10–20" sistemos EEG elektrodų išdėstymo pozicijos [12]. Inion - žemiausia pakaušio dalis, Nasion - žemiausia kaktos dalis. Lyginiais skaičiais žymima dešinė galvos pusė, nelyginiais kairė. Raidėmis F, T, C, P ir O žymimos priekinės, smilkinės, centrinės, išorinės ir pakaušio galvos sritys. Raidė A žymi ausų vietą.

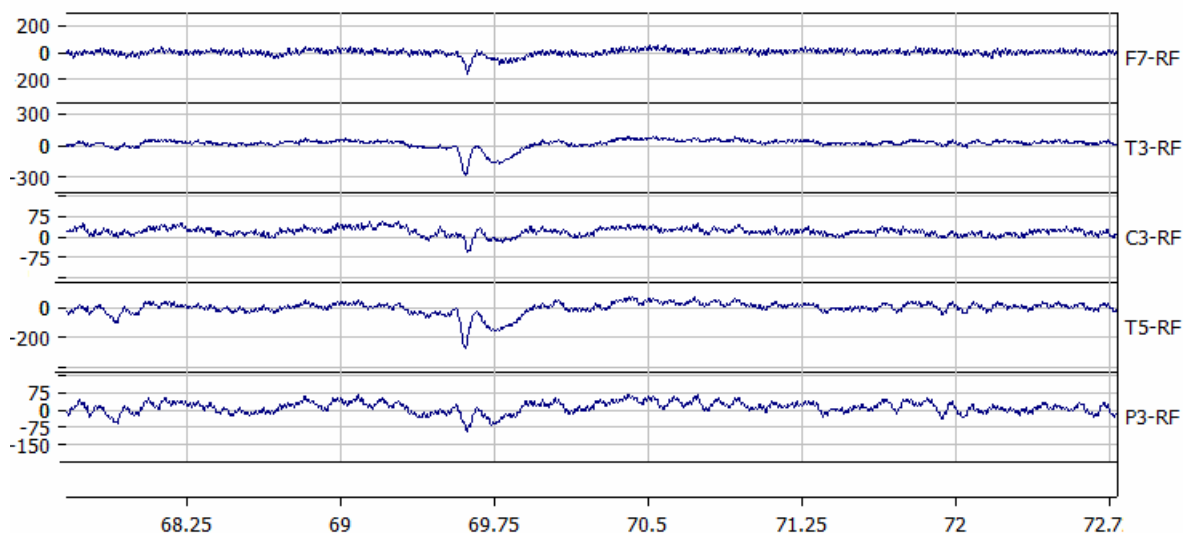
1.2. Smegenų veikla ir epileptoforminių pikų požymiai

Normaliai smegenų veiklai, būdingi įvairūs smegenų ritmai. Žmogui net ir esant ramybės būsenoje, smegenys yra aktyvios, bei vykdo įvairias veiklas: pradedant nuo sensorinės informacijos apdorojimo, baigiant planavimu. Dėl šios priežasties EEG gali būti atliekamos tiek budrumo būsenoje, tiek paciento miego metu. Epilepsijos aptikimo atveju, abi EEG atlikimo būsenos turi vienodą vertę. Suaugusiems pacientams EEG dažniausiai atliekama budrumo metu, negalintiems išbūti ramiai vaikams - miego metu. Užrašius EEG, gydytojas vizualiai ieško, neįprastos smegenų veiklos, suaktyvėjimo atvejų - epileptoforminių pikų. Kartais kai kurie signalų artefaktai yra vizualiai panašūs į epileptoforminius pikus, todėl svarbu suformuluoti epileptoforminio piko morfologines savybes.

EEG signalo pikui, kad jis būtų laikomas epileptoforminiu keliami tam tikri morfologiniai reikalavimai [7]:

- viso piko trukmė turi būti tarp 40 – 200 ms.
- pikas turi iškilti bent du kartus virš foninio smegenų aktyvumo.
- pikas turi atsikartoti bent dvejuose kaimyniniuose kanaluose.
- dažniausiai aptinkamas C4/C3 ir C5/C6 kanaluose.

Epileptoforminiai pikai nepriklausomai vienas nuo kito, gali būti aptinkami skirtingose galvos srityse. Pasižymi aukšta amplitude (100 – 300 μV), dažniausiai juos lydi iškrovos banga, taip pat gali pasirodyti tiek pavieniai, tiek pikų grupėse.



2 pav. Vaizduojamas epileptoforminis pikas ties 69.7 sekunde, matomas tiek židinyje (T5 kanalas), tiek jo kaimyniniuose kanaluose. Signalas matuojamas mikrovoltais per milisekundę.

2. EEG pikų aptikimas

2.1. Susijusių darbų apžvalga

Įvedus paieškos raktinį žodį *seizure detection* Google Mokslinčiujė (Google Scholar), gausime daugiau 200 000 mokslinių straipsnių šia tema, tai įrodo, jog elektroencefalogramų analizės automatizavimas vis dar yra aktuali sritis [2]. Tačiau pilnai automatizuoti pikų aptikimą yra sudėtinga dėl daugelio priežasčių. Visų pirma, reikalingas glaudus bendradarbiavimas su neurologais, nes kiekvienam pacientui būdingi specifiniai pikai ir galutinei diagnozei, dažnu atveju reikalinga bent trijų skirtingų gydytojų nuomonė pikams nustatyti [15]. Taip pat epileptoforminių pikų detekcijos algoritmų testavimui ar apmokymui (naudojant specifinius algoritmus) reikalingos skirtingos EEG signalų bazės. Jos turi būti pakankamai didelės, kad galėtų reprezentuoti visą populiaciją.

Šiuo metu egzituojančius automatinės analizės algoritmus, galima skirstyti pagal jų naudojamus matematinius principus. Pagrindinės metodų grupės yra: šablono atitikties radimo, parametrinių metodų, mimetrinės analizės, bangų transformacijos ir dirbtinių neuroninių tinklų [6].

Pastarosios dvi kategorijos yra bene plačiausiai plėtojamoms bei pasižymi dideliu tikslumu [5], tačiau dirbtinių neuroninių tinklų algoritmų apmokymui reikalingos didelės duomenų imtys, taip pat duomenys apmokymui turi būti geros kokybės, norint išgauti kuo didesnę tikslumą [8].

Lyginti algoritmus tarpusavyje yra sudėtinga ne tik dėl to, jog jie ramiasi skirtinga teorija, metrikomis, bet ir dėl to, kad jie naudoja skirtingus, specifinius jų atvejui, EEG signalų rinkinius [6].

Šiame darbe bus kalbama apie du epileptoforminių pikų detekcijai skirtus algoritmus. Pirmasis, kuris remiasi morfologinių operacijų ir filtrų teorija [16], tačiau galėtų būti kategorizuojamas ir visiškai atskirai, vadinant jį atvirkštiniu šablono atitikties algoritmu [10]. Ir antrasis algoritmas, kuris epileptoforminius pikus aptinka signalą interpoliuojant gabalais tiesine funkcija.

2.2. EEG pikų aptikimo algoritmas, remiantis morfologiniais filtrais

Epileptoforminių pikų detekcijos automatizavimui pasirinktas algoritmas remiasi morfologinių filtrų teorija. Jo vykdymo metu, atliekamos įvairios morfologinės operacijos, kurios nufiltruoja smegenų foninę veiklą ir išryškina ligos pikus. Taip pat algoritmas pasižymi gana aukštu aptikimo tikslumu - 91.62 % [16].

2.2.1. Morfologinių filtrų teorija

Norint įgyvendinti morfologinių filtrų algoritmą, pirmiausia reikia apibrėžti elementariausias morfologines operacijas.

Tiriamas EEG signalas - tai laiko eilutė $f(t)$. Struktūrinis elementas - tai funkcija $g(t)$, o jo atspindys apibrėžiamas: $g^s(t) = g(-t)$, kur $D - f(t)$ apibrėžimo sritis.

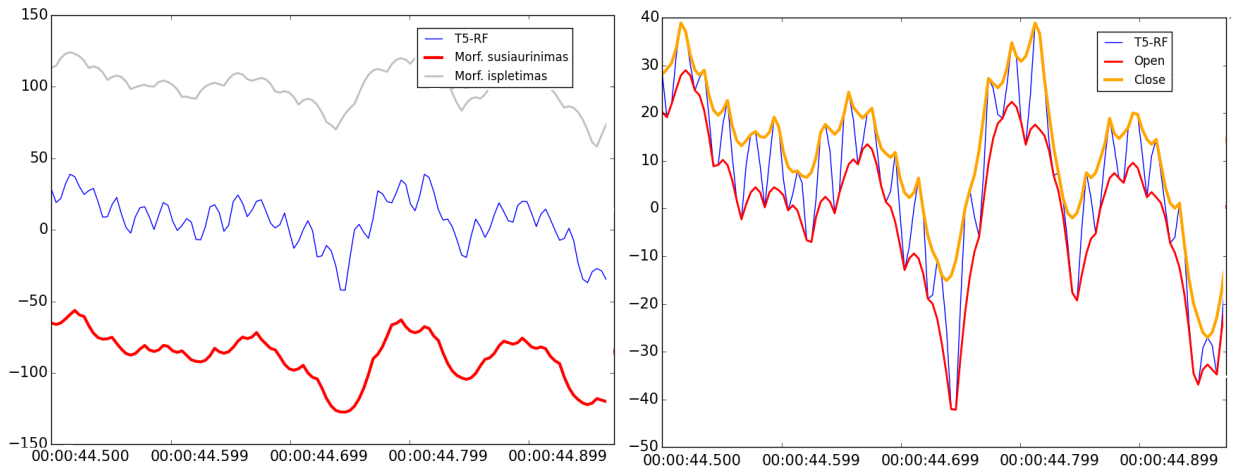
Morfologiniuose filtruose naudojamos dvi pagrindinės matematinės operacijos: morfologinis susiaurinimas ir išplėtimas.

Morfologinis susiaurinimas:

$$(f \ominus g^s)(t) = \min_{\tau \in D} f(\tau) - g(-(t - \tau)) \quad (2.1)$$

Morfologinis išplėtimas:

$$(f \oplus g^s)(t) = \max_{\tau \in D} f(\tau) + g(-(t - \tau)) \quad (2.2)$$



3 pav. Kairėje vaizduojamas originalus signalas bei atliktos pagrindinės elementariosios morfologinės operacijos : morfologinis susiaurinimas ir išplėtimas. Dešinėje originalus signalas ir atliktos morfologinio atidarymo bei uždarymo operacijos. Morfologinės operacijos buvo atliktos su parabolės formos struktūriniu elementu, kurio plotis 19 signalo taškų (0.074 s), aukštis apie 90 mikrovoltų. Signalai matuojami mikrovoltais per milisekundę.

3 paveikslėlyje kairėje vaizduojamas pagrindinių elementariųjų morfologinių operacijų veikimas realiame paciento signale. Struktūriniu elementu pasirinkta parabolė, sudaryta iš 19 taškų, kas atitinka 0.074 s ir yra apie 90 mikrovoltų aukščio. Morfologinio susiaurinimo metu, pasirinktu struktūriniu elementu judama per signalą ir ieškoma minimali reikšmė. Kadangi pasirinkta parabolė buvo sąlyginai neaukšta, tačiau gana plati, po morfologinio filtro gautajame signale pikai yra nebe tokie aštrūs ir paslukti parabolės aukščiu žemyn. Morfologinis išplėtimas veikia priešingai: gautasis signalas yra pakeltas struktūrinio elemento aukščiu į viršų.

Žinant šias pagrindines morfologines operacijas, galima konstruoti signalo ir jo struktūrinio elemento kompozicijas, apibrėžiamas kaip morfologinis uždarymas ir atidarymas.

Morfologinis atidarymas:

$$(f \circ g)(t) = [(f \ominus g^s) \oplus g](t) \quad (2.3)$$

Morfologinis uždarymas:

$$(f \bullet g)(t) = [(f \oplus g^s) \ominus g](t) \quad (2.4)$$

Morfologinis atidarymas signalo $f(t)$ pikus suapvalina iš apačios, sumažindamas jų reikšmę, tuo tarpu morfologinis uždarymas pakelia signalo reikšmes (3 pav. dešinėje). Vadinasi, morfologinis atidarymas ir uždarymas gali aptikti tik arba teigiamos, arba neigiamos reikšmės pikus. Kadangi EEG pikai gali būti tiek teigiami, tiek neigiami, todėl būtina apibrėžti lankstenes operacijas.

Morfologinis uždarymas - atidarymas:

$$CO(f(t)) = f(t) \bullet g_1(t) \circ g_2(t), \quad (2.5)$$

Morfologinis atidarymas - uždarymas:

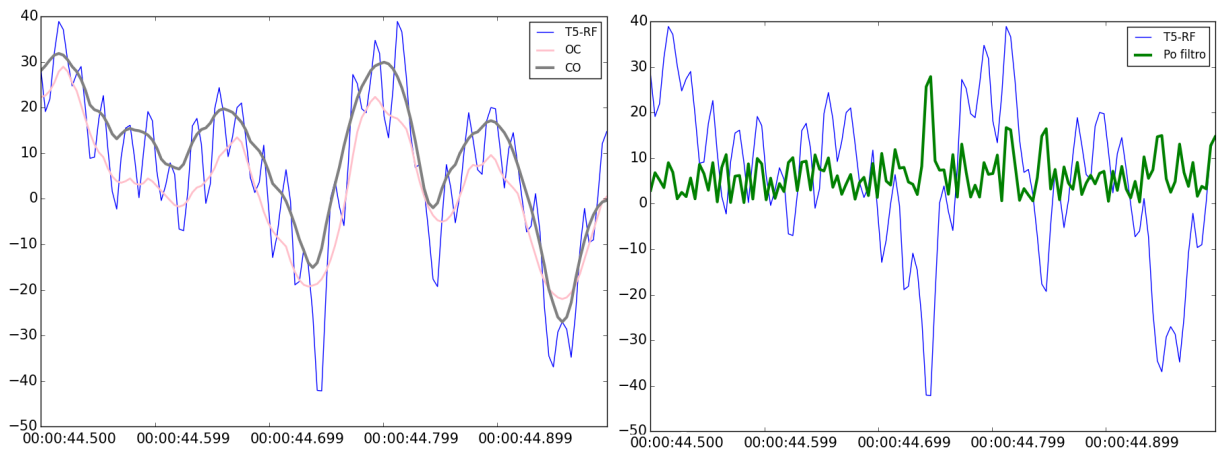
$$OC(f(t)) = f(t) \circ g_1(t) \bullet g_2(t), \quad (2.6)$$

$g_1(t)$ ir $g_2(t)$ - struktūriniai elementai. Šios abi aprašytos operacijos, deformuoja signalą: atidarymo - uždarymo operacija sumažina signalo reikšmes, o uždarymo - atidarymo - pakelia (4 pav. kairėje). Tokia deformacija, gali sukelti neatitikimų pikų aptikime, todėl reikia sukonstruoti filtrą, kuris būtų minėtųjų operacijų vidurkis:

$$OCCO(f(t)) = \frac{(OC(f(t)) + CO(f(t)))}{2} \quad (2.7)$$

Signalas $x(t)$, gautas pritaikius morfologinius filtrus, aprašomas formule [7]:

$$x(t) = f(t) - OCCO(t) \quad (2.8)$$



4 pav. Kairėje vaizduojamas originalus signalas bei atlikti OC ir CO morfologiniai filtrai. Dešinėje originalus signalas ir galutinis rezultatas, aprašytas 2.8 formule. Morfologinės operacijos buvo atliktos su parabolės formos struktūriniu elementu, kurio plotis 19 signalo taškų (0.074 s), aukštis apie 90 mikrovoltų. Signalai matuojami mikrovoltais per milisekundę.

Paveikslėlio viršuje dešinėje, matomas originalus signalas ir skirtumas tarp originalaus signalo, bei nufiltruoto OCCO filtru (2.7 formulė). Grafike atvaizduojamas ne pats galutinis signalas po šio morfologinio filtro panaudojimo, bet jo modulis. Taip gaunamos tik teigiamas galutinės reikšmės. Kadangi epileptoforminiai pikai gali būti tiek teigiami, tiek neigiami, jų detekcijai reiktų nustatyti dvi ribas: vieną neigiamiems pikams, kitą - teigiamiems. Dėl šios priežasties, yra paranku turėti tik teigiamų reikšmių signalą ir taip apsisistoti ties viena detekcijos riba.

2.2.2. Struktūrinio elemento parinkimas

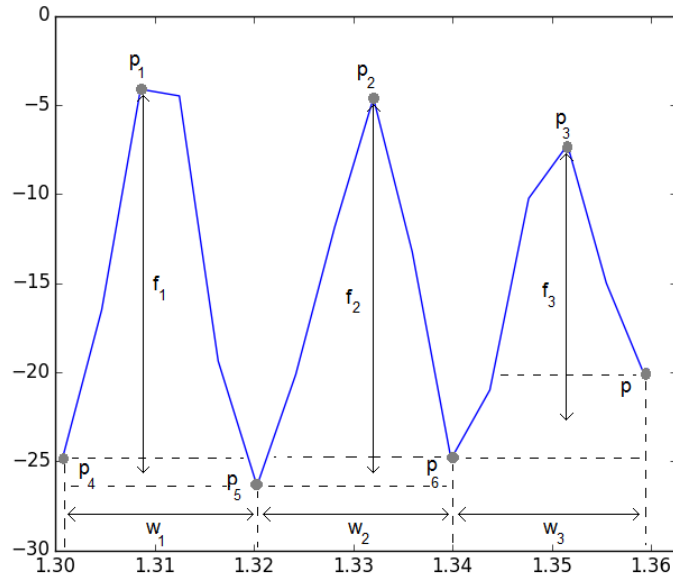
Morfologinių filtrų teorija dažniausiai yra taikoma skaitmeniniam vaizdų apdorojimui, todėl struktūriniais elementais yra pasirenkami įvairių formų objektai, kuriais norima apdoroti paveikslėlį. Encefalogramų apdorojimo atveju, struktūrinius elementus reikia parinkti taip, kad morfologinio filtravimo metu, būtų nufiltruotas tik foninis smegenų aktyvumas, paliekant epileptoforminius pikus. Smegenų veiklai būdingi periodiniai ritmai, todėl foniniam smegenų aktyvumui aprašyti pasirinkta parabolė:

$$g_i(t) = a_i t^2 + b_i, i = 1, 2 \quad (2.9)$$

Struktūrinio elemento parinkimo tikslas - tai parabolės sukonstravimas, kuri geriausiai nufiltruotų foninę smegenų veiklą. Kadangi signale esančių parbolių amplitudžių aukščiai dažniausiai skiriasi kiekvienoje EEG atkarpoje, todėl koeficientai a ir b , turi kisti kartu su fonine smegenų veikla. Taip bus užtikrinamas labiausiai tinkamos parabolės radimas. Optimaliam koeficientų parinkimui yra apskaičiuojami signalo arkų pločiai W , bei arkų ilgiai $|f|$.

W - distancijos tarp dviejų gretimų signalo minimumų arba maksimumų,

$|f|$ - signalo aukščių amplitudės.



5 pav. Signalo arkų pločių W , bei arkų ilgių $|f|$ apskaičiavimo pavyzdys realiaame signale. Kadangi signalo minimumai p_5 ir p_6 nėra viename aukštyje, arkos aukštis f_2 yra apskaičiuojamas, kaip abiejų šių viršūnų aukščių vidurkis. Signalas matuojamas mikrovoltais per milisekundę.

Žinant W ir $|f|$, galima abibrėžti a ir b :

$$a_1 = \frac{\text{median}|f|}{0.5\text{median}|W|}, \quad a_2 = \frac{\text{median}|f|}{1.5\text{median}|W|} \quad (2.10)$$

$$b_1 = b_2 = \text{median}|f| \quad (2.11)$$

Apskaičiavus koeficientus a ir b , gaunamos dvi parabolės. Jos yra vienodo pločio, tačiau skiriasi savo aukščiu. Gautieji struktūriniai elementai dalyvauja signalo filtravime: elementariosiomis morfologinių filtrų operacijomis, filtruojamos parabolės, kurios apibūdina foninę smegenų veiklą. Taip pat dėl struktūrinių elementų aukščių skirtumo, labiau yra išryškunami epileptoforniniai pikai.

Struktūrinių parbolių aukščių skirtumo tyrimas yra pateikiamas 4.2 ir 4.3 skyreliuose.

2.2.3. EEG signalo tiriamos atkarpos ilgio parinkimas

Dėl nuolat kintančios smegenų veiklos, struktūrinis elementas turi būti pastoviai perskaičiuojamas. Parinkti tinkamą analizei atkarpos ilgį, kuriam reiktų perskaičiuoti struktūrinį elementą iš

naujo, yra gana sudėtinga, nes EEG atkarpa turi būti pakankamai ilga, jog galėtume, joje aptikti epileptiforminį piką, tačiau kartu ir pakankamai trumpa, kad smegenų vyraujantys ritmai nepasikeistų. Žinant jog, maksimali epileptiforminio piko trukmė yra 200 ms, atkarpa analizei reikia pasirinkti ilgesnę, kad apskaičiuotos foninės veiklos parabolės atspindėtų ne tik patį piką, bet ir smegenų veiklą iki jo pradžios.

Signalų kanalas	Laiko tarpas	Parabolių aukščiu med.	Parabolių pločiu med.
T5-RF	0 min – 20 min	16	12
T5-RF	0 min – 10 min	16	12
T5-RF	0 min – 5 min	18	13
T5-RF	0 min – 1 min	18	13
T5-RF	0 – 30 s	17	12
T5-RF	0 – 15 s	15	12
T5-RF	0 – 10 s	17	13
T5-RF	0 – 5 s	18	13
T5-RF	0 – 3 s	18	13

1 lentelė. Paciento smegenų veiklos parabolių arkų pločiai ir aukščiai realiame signale. Aukštis matuojamas mikrovoltais, plotis vienetais, atitinkančiais 0,0039 s. Signalas prieš tai buvo nufiltruotas Furjė filtru.

Literatūroje yra minima, jog optimali signalo atkarpa yra tarp 4 ir 6 sekundžių [10]. Iš pateiktos 1 lentelės duomenų, galima patvirtinti tokį atkarpos pasirinkimą: jei algoritme būtų naudojamos foninės smegenų veiklos parabolių medianos apskaičiuotos visam signalo laikui, jos idealiai neatspindėtų tiriamos trumpos epileptiforminio piko atkarpos. Taip būtų susidaromi struktūriniai elementai, kurie įneštų neatitikimų pikų detekcijoje. Kaip matome, signalo duomenys lentelėje tarp 3 ir 5 sekundžių nesiskiria, todėl tolimesniuose šio darbo etapuose buvo naudojama 5 sekundžių atkarpa struktūrinio elemento perskaičiavimui.

2.2.4. Struktūrinio elemento ilgio parinkimas

Kaip jau yra žinoma, struktūrinis elementas yra parabolė. Struktūrinio elemento ilgis - tai taškų skaičius, iš kurių yra sudaryta parabolė. Struktūrinis elementas kinta, jei yra keičiami koeficientai a ir b (2.9, 2.10 ir 2.11 formulės). Tačiau tai nedaro įtakos parabolės pločiui. Koeficientai a ir b aukština struktūrinę parabolę, tuo tarpu jos plotis yra iš anksto nustatytomas formule [10]:

$$S_L = 4median|W|, \quad (2.12)$$

kur S_L - Struktūrinio elemento ilgis, $|W|$ - signalo arkų pločiai.

Parabolės plotis yra vienas esminių veiksmų pikų filtravime. Jei parabolė bus itin siaura, sudaryta vos iš kelių reikšmių, signalas bus per daug nufiltruotas, jei ji labai plati, sudaryta iš daugybės taškų, foninė veikla išliks. 2.12 formulėje aprašyto struktūrinio elemento ilgio nustatymas buvo patvirtintas atliekant detalesnius tyrimus su realių pacientų duomenimis. Išsamesnė analizė aprašyta 4.1 skyrelyje.

2.2.5. Pikų detekcijos riba

Po morfologinių filtrų panaudojimo, gautasis rezultatas (2.8 formulė) - tai skirtumas tarp originalaus paciento signalo ir signalo, nufiltruoto morfologiniais filtrais. Šis skirtumas, tai mažesnės amplitudės už originalų signalas, kuriame smegenų foniniai ritmai nebe tokie ryškūs, o epileptoforminiai pikai išsiskiria savo aukščiu.

Detekcijos riba - tai nustatyta reikšmė, virš kurios aptinkamos signalo viršūnės jau yra laikomos pikais. Kadangi epileptoforminiai pikai gali būti tiek teigiami, tiek neigiami, reikalingos dvi detekcijos ribos: teigiamoms ir neigiamoms reikšmėms. Siekiant supaprasti detekcijos ribos radimą, naudojamas nufiltruoto signalo modulis $|x(t)|$, taip nustatydamas tik teigiamus turimo signalo rezultatus. Apskaičiavus visų gautojo signalo viršūnių amplitudes, išvedamas jų vidurkis:

$$\bar{x}(t) = \frac{1}{n} \sum_{i=1}^n |x(t)|, \quad (2.13)$$

n - turimų viršūnių skaičius.

Tuomet epileptoforminių pikų aptikimui naudojama detekcijos riba L randama:

$$L = d * \bar{x}(t), \quad (2.14)$$

kur d - detekcijos ribos koeficientas.

Kitų autorių darbuose [7], siūloma naudoti nufiltruoto signalo modulių medianos reikšmę, tačiau atlikus analizę su realiais pacientų duomenimis, buvo likta prie 2.14 formulės.

Detekcijos ribos koeficientas d - tai konstanta, kuri dėl skirtingų pacientų savybių, kiekvienai encefalogramai gali būti skirtinga. Vienas iš tikslų šiame darbe yra šio koeficiento radimas. Šie tikslinimo darbai aprašyti 4.4 skyrelyje.

2.3. EEG pikų aptikimo algoritmas, remiantis interpoliavimu gabalais tiesine funkcija

Antrasis epileptoforminių pikų detekcijos automatizavimui pasirinktas algoritmas remiasi interpoliavimu gabalais tiesine funkcija. Jo vykdymo metu, tiesėmis interpoliuojamos duomenų atkarpos. Kadangi epileptoforminiai pikai yra paciento galvos smegenų impulsų suaktyvėjimas, jie pasižymi staigiais signalo pokyčiais encefalogramose: signalas staigiai kyla, bei leidžiasi. Atvaizduojant šiuos pokyčius tiesėmis, pasirinktojoje atkarpoje nubrėžiamos 5 tiesės taip atvaizduojant epileptoforminio piko šablona. Todėl gautajame naujame signale po interpoliavimo gabalais tiesine funkcija, pikai išlieka, o smegenų foninė veikla "išnyksta" ją atvaizduojant tiese. Taip gaunamas galimų epileptoforminių pikų sąrašas. Vėliau juos atfiltruojant pagal tam tikrus tiesių, kuriomis interpoliuojame, koeficientų požymius, detektuojami tikrieji pikai.

2.3.1. Gabalais tiesinė funkcija

Matematikoje, gabalais tiesinė funkcija (angl. piecewise linear function) vadinama funkcija, sudaryta iš realiųjų skaičių konstantų arba iš tiesių segmentų, aprašančių skirtingas funkcijos atkarpas, kurias jungia tam tikros sąlygos [13]. Tarkime turime taškus x_0, x_1, \dots, x_n , kurie tenkina sąlygą:

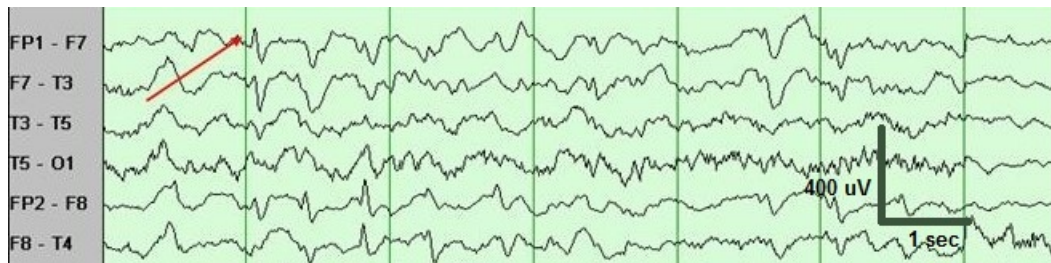
$$x_0 < x_1 < \dots < x_n \quad (2.15)$$

Tada gabalais tiesinė funkcija $f(x)$ - tai tiesių rinkinys, kuris kiekviename intervalo $[x_{i-1}, x_i]$ taške yra lygus $k = 1$ laipsnio polinomui:

$$f(x) = \begin{cases} f_1(x) = b_1 + k_1x, & x \in [x_0, x_1], \\ f_2(x) = b_2 + k_2x, & x \in [x_1, x_2], \\ \vdots \\ f_n(x) = b_n + k_nx, & x \in [x_{n-1}, x_n]. \end{cases} \quad (2.16)$$

2.3.2. Teorinis algoritmo aprašymas

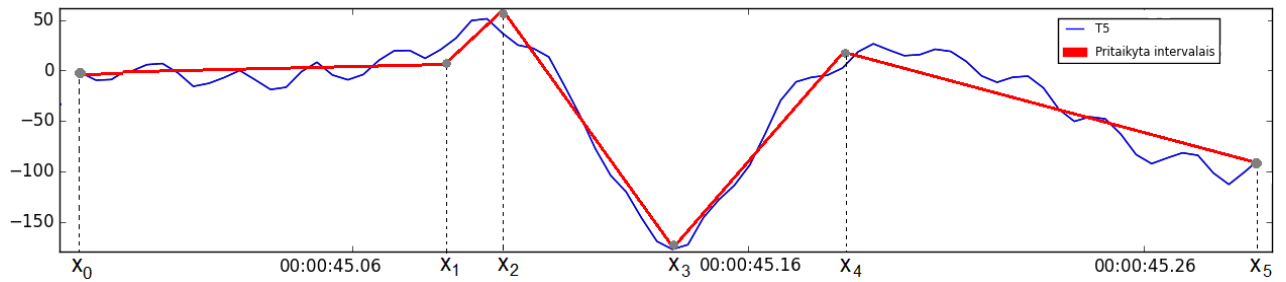
Epileptiforminiai pikai skirtingiems pacientams gali turėti skirtingus parametrus: aukštis, plotis, trukmė, kiekis. Tačiau bendrinio atveju, pikų forma yra žinoma [1]. Kaip matome paveikslėlyje žemiau (6 pav.), epileptiforminis pikas, tai staigus signalo pakilimas, vėliau kritimas ir vėl pakilimas iki normalios smegenų veiklos.



6 pav. Matomas epileptiforminio piko pavyzdys iš internetinio žurnalo Medscape [1].

Iš 2.16 formulės žinoma, jog gabalais tiesinę funkciją sudaro atskiros tiesės, skirtinguose laiko intervaluose. Piko atveju, pirmoji tiesė atitiktų normalią smegenų veiklą nuo tiriamos atkarpos pradžios x_0 iki piko pradžios x_1 . Antroji atvaizduotų piko kilimą nuo piko pradžios x_1 iki aukščiausio piko taško x_2 . Trečioji - epileptiforminio piko ledimasis nuo aukščiausio taško x_2 iki žemiausiojo x_3 . Ketvirtoji atitiktų piko grįžimą iki normalios smegenų veiklos nuo žemiausio taško x_3 iki x_4 . Penktoji tiesė - normali smegenų veikla po įvykusio suaktyvėjimo intervale $[x_4, x_5]$ (7 pav.). Taigi naudojamam algoritmui užtenka gabalais tiesinės funkcijos, kurią sudaro penkios tiesės:

$$f(x) = \begin{cases} f_1(x) = b_1 + k_1x, & x \in [x_0, x_1], \\ f_2(x) = b_2 + k_2x, & x \in [x_1, x_2], \\ f_3(x) = b_3 + k_3x, & x \in [x_2, x_3], \\ f_4(x) = b_4 + k_4x, & x \in [x_3, x_4], \\ f_5(x) = b_5 + k_5x, & x \in [x_4, x_5]. \end{cases} \quad (2.17)$$



7 pav. Gabalais tiesinės funkcijos pritaikymas viename iš encefalogramos kanalų. Vaizduojamoje atkarpoje išvedamos 5 tiesės, artimiausios pačiam signalui. $x_0, x_1, x_2, x_3, x_4, x_5$, tai tiesių pradžių ir pabaigų taškai. Signalas matuojamas mikrovoltais per milisekundę.

2.3.3. Gabalais tiesinės funkcijos skaičiavimas

Nagrinėjama gabalais tiesinė funkcija - tai tiesių rinkinys, sudarytas iš 5 tiesių. Kiekviena tiesė turi skirtingus krypties koeficientus k ir narij b , nurodantį tašką, kuriame tiesė kerta $f(x)$ ašį. Taip pat yra žinomos visos laiko reikšmės nuo pradžios taško x_0 iki pabaigos taško x_5 bei visos signalo reikšmės bet kuriuo pasirinktu laiko momentu $f(x_0), \dots, f(x_5)$. Sprendžiamas uždavinys - tai parametų $k_j, b_j, x_1, x_2, x_3, x_4$ radimas, kurie būtų optimaliausi, mažiausių kvadratų metodo atžvilgiu, kai $j \in [1, 5]$.

2.3.4. Mažiausių kvadratų metodo naudojimas tiesių radimui

Ieškomų tiesių per duotuosius taškus radimui naudojamas mažiausių kvadratų metodas. Šis metodas dažniausiai yra naudojamas lentele išreikštų funkcijų aproksimavimui. Turint taškus x_0, x_1, \dots, x_n ir žinant funkcijos $f(x_i)$ reikšmes šiuose taškuose, kai $i \in [0, n]$, funkciją $f(x)$ galima vadinti empirine funkcija bei reikalauti, kad atstumai tarp empirinės funkcijos reikšmių ir eksperimento taškų būtų minimalūs.

Mažiausių kvadratų metodas apskaičiuoja tinkamiausius tiesės parametus minimizuojant nuokrypių sumą. Funkcijos nuokrypiu yra laikomas skirtumas tarp žinomos funkcijos reikšmės ir apskaičiuotos reikšmės [14]. Mažiausių kvadratų metodo reikalavimas yra toks: ieškomi parametrai turi būti tokie, kad nuokrypių kvadratų suma būtų minimali:

$$S = \sum_{i=1}^n (f_i - y_i)^2, \quad (2.18)$$

kur y_i - empirinė funkcija, f_i - apskaičiuota reikšmė

Ieškoma suma, kuri būtų kiek įmanoma mažiausia, tai galima padaryti ieškant funkcijos minimumo.

Pirmos eilės išvestinės prilyginimas nuliui nurodo funkcijos ekstremumus. Antros eilės išvestinė indikuoja, ar ekstremumo taškas yra minimumas ar maksimumas [4]. Jei antroji išvestinė yra teigiama konstanta, tai pirmosios išvestinės prilyginimas nuliui turės vienintėlį sprendinį ir funkcija su šiuo sprendiniu keičia ženklą iš neigiamos į teigiamą - taip gaunamas funkcijos minimumas.

Formulės išvestinių skaičiavimui:

$$\frac{ds}{db_t} = 2 \sum_{i=1}^n (f_i - y_i) \frac{df_i}{db_t} \quad (2.19)$$

$$\frac{d^2s}{db_t^2} = 2 \sum_{i=1}^n \left((f_i - y_i) \frac{d^2s}{db_t^2} + \left(\frac{df_i}{db_t} \right)^2 \right) \quad (2.20)$$

Kur t - ieškomų parametrų indeksai, $t = 1, \dots, n$. Ieškomi parametrai yra randami pirmąją išvestinę prisilyginus nuliui ir išsprendus lygčių sistemą :

$$\sum_{i=1}^n (f_i - y_i) \frac{df_i}{db_t} = 0 \quad (2.21)$$

2.3.5. Tiesių koeficientų reikšmė

Gabalais tiesinę funkciją (2.17 formulė), sudaro 5 tiesės, kurių kiekviena turi koeficientus k_i ir b_i , kur $i = 1, \dots, 5$. Tiesės krypties koeficientas k – skaičius, apibūdinantis tiesės kryptį ir nuolydį [3]. Matematiškai tiesės krypties koeficientą k galima užrašyti kaip:

$$k = \frac{y_2 - y_1}{x_2 - x_1} \quad (2.22)$$

Norint sužinoti krypties koeficientą, yra skaičiuojamas santykis tarp dviejų skirtingų tiesės taškų y koordinatų su x koordinatų skirtumu. x koordinatų skirtumas - tai ilgių skirtumas, o y koordinatų skirtumas - tai tiesę aprašančių dviejų taškų aukščių skirtumas.

Tiesė yra didėjanti, jei, eina iš kairės į dešinę, kyla į viršų. Tokiu atveju, krypties koeficientas k teigiamas : $k > 0$. Tiesė vadinama mažėjančia, kai leidžiasi žemyn. Krypties koeficientas neigiamas : $k < 0$. Tiesės nuolydis - tai tiesės statusas, matuojamas krypties koeficiento absoliučiaja reikšme. Kuo didesnė absoliučioji reikšmė, tuo tiesė statesnė.

Žinant šiuos tiesių koeficientų apibrėžimus, yra formuluojamos keletas savybių, kurios yra būdingos epileptoforminiams pikams. Epileptoforminiai pikai - tai stiprus smegenų suaktyvėjimas ir staigus funkcijos, aprašančios smegenų veiklą, didėjimas arba mažėjimas. Dėl šios priežasties tiesių nuolydžiai piko atveju turėtų būti pakankamai didelės reikšmės ir smarkiai skirtis nuo kitų tiesės aprašančių krypties koeficientų:

$$||k_1| - |k_2|| > const. \quad (2.23)$$

7 paveikslėlyje matoma, jog pirmoji tiesė f_1 ir antroji f_2 yra skirtingų statusų, todėl tarp jų yra ryškus modulių įverčių skirtumas. Kadangi nesvarbi tvarka iš kurios absoliučiosios vertės atimsime kurią, 2.23 formulėje yra naudojamas dar vienas modulio ženklas. Tolimesniuose šio darbo skyriuose yra atliekamas šios konstantos analizė ir pateikiamos konkrečios jos reikšmės.

Kita epileptoforminiams pikams būdinga savybė, tai tiesių didėjimas ir mažėjimas. Todėl koeficientų k sandauga tarp f_1 ir f_2 yra neigimas skaičius:

$$|k_1| * |k_2| < 0 \quad (2.24)$$

Teigiama šių dviejų koeficientų sandauga žymėtų dviejų tiesių kylimą arba leidimąsi, todėl tai nebūtų išryškintas epileptoforminis pikas, o tiesiog dvi smegenų veiklą aprašančios tiesės su vienoda kryptimi.

Po interpoliavimo gabalais tiesine funkcija, gaunamas signalas, kuriame vis dar ryški foninė smegenų veikla. Pritaikius šias dvi aprašytas tiesių koeficientų savybes, signale detekcijai išlieka tik epileptoforminiai pikai.

3. Pagalbiniai įrankiai algoritmų įgyvendinimui

3.1. Testinių duomenų bazė

Duomenys, skirti analizei, buvo gauti iš Vaikų ligoninės, viešosios įstaigos Vilniaus universiteto ligoninės Santariškių klinikų filialo. Encefalogramos, skirtos nagrinėjimui, yra vaikų, kuriems nustatyta Rolando epilepsija. Viso duomenų bazę sudaro 21 encefalograma, iš kurių 4 encefalogramos yra apie 4 minučių ilgio, 1 apie 36 minučių ilgio. Likusiose signalų trukmė apie 20 minučių. Kiekvieną encefalogramą sudaro 21 kanalas. Ši duomenų bazė turėjo detalius aprašymus, kuriame iš kanalų yra aptikti ligą žymintys pikai, koks nustatytas jų skaičius, kuriose kaimyniniuose kanaluose yra matomas atsikartojimas. Taip pat buvo žinoma pacientų lytis, tikslus vaikų amžius nebuvo pateikiamas.

Visa suteikta informacija buvo naudinga morfologinių filtrų algoritmo parametrų nustatymuose: žinant tikslų pikų skaičių, įvairiai keičiant parametrų reikšmes, buvo nustatomos optimaliausios algoritmo skaičiavimo galimybės. Gabalais tiesinės funkcijos pritaikymui, tikslus pikų skaičiaus žinojimas buvo naudingas tinkamo laiko intervalo parinkimui bei žingsnio, kuriuo algoritmas turi būti vykdomas nustatymui.

Paciento lytis	Maks. pikų skaičius	Min. pikų skaičius	Vidutinis pikų skaičius
14 berniukų	1060	22	313
7 mergaitės	398	2	125
viso 21 pacientas	1060	2	251

2 lentelė. Turimos duomenų bazės pikų skaičiai pacientų encefalogramose.

3.1.1. Programoje naudotos bibliotekos

Encefalogramų analizei skirtų morfologinių filtrų algoritmo realizavimas buvo įgyvendintas Python atvirojo kodo programavimo kalba. Ši kalba buvo pasirinkta dėl gausybės išbaigtų matematinių bibliotekų, bei sintaksės paprastumo. Šių programavimo kalbos savybių dėka, morfologinių filtrų ir interpoliavimo gabalais tiesine funkcija įgyvendinimas tampa paprastesnis, neapkraunantis kodo įvairiais rankiniais skaičiavimais ir lengviau suprantamas vartotojui. Algoritmo realizacijai palengvinti buvo naudojamos įvairios Python kalboje esančios bibliotekos:

- SciPy - įrankis skirtas įvairiems techniniams, matematiniais skaičiavimams. Bibliotekoje gausu paketų skirtų tiesinei algebrai, integravimui, signalų ir paveikslėlių apdorojimui. Morfologinių filtrų operacijoms buvo naudotos funkcijos iš SciPy bibliotekos, ndimage paketo.
- NumPy - įvairiems skaičiavimams atlikti ir darbui su skaičių masyvais skirta biblioteka. Programoje ji ne kartą gelbėjo skaičiuojant medianą arba rikiuojant reikšmes.
- EEGtools - įrankis encefalogramų nusiskaitymui ir patogiam tolimesniam darbui su signalais.
- matplotlib - biblioteka, skirta grafiniam rezultatų pateikimui.

- PyQt4 - grafinei sąsajai kurti bei palaikyti skirta biblioteka. Ją sudiegus, kartu buvo gautas įrankis grafinei sąsajai kurti - Qt Designer. Šio įrankio pagalba buvo lengvai sukurta GUI forma, algoritmų koeficientų tikslinimui.
- glob - modulis, skirtas duomenų prieigai. Jo pagalba buvo grąžinamos tiriamų encefalogramų failų nuorodos iš aplanko kompiuteryje.
- csv - modulis duomenų nusiskaitymui ir kėlimui CSV formatu. Programoje gauti koeficientai buvo išvedami į failą CSV formatu rankinei duomenų peržiūrai.

3.2. Naudoti bei sukurti įrankiai analizei

Vienas iš šio magistro baigiamojo darbo tikslų - algoritmuose naudojamų skaičiavimo koeficientų tikslinimas.

Morfologinių filtrų algoritmo koeficientų tikslinimui ir detekcijos ribos radimui palengvinti buvo sukurta grafinė sąsaja (angl. graphical user interface (GUI)), kurios pagalba buvo lengvai vizualiai matoma vieno koeficiento priklausomybė nuo kito (8 pav.). Grafinė sąsaja buvo sukurta naudojant Python GUI įrankį Qt Designer.



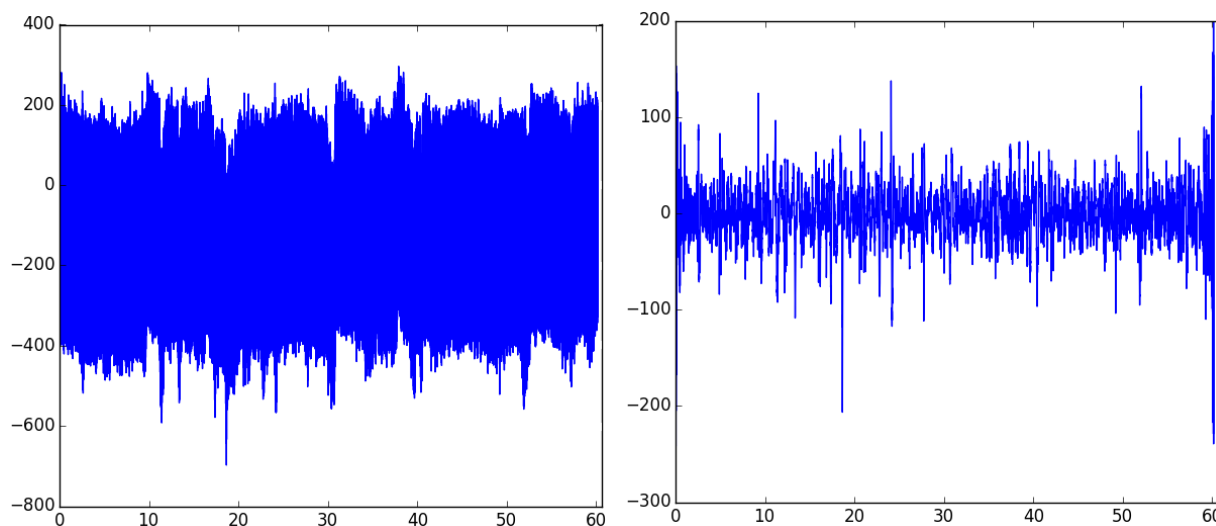
8 pav. Sukurtoje grafinės sąsajos formoje, yra galimybė pasirinkti norimą signalą, jo tiriamą kanalą, bei keičiant H (aukštis) ir W (plotis), stebėti kaip kinta pasirinkto signalo grafikas, bei detektuojamų pikų skaičius

Grafinėje sąsajoje yra galimybė peržiūrėti atliktas morfologines operacijas grafiškai. Pažymėjus norimą filtrą varnele, jis atvaizduojamas greta originalaus signalo. Tai palengvino struktūrinio elemento koeficientų nustatymą: matant vizualiai kaip keičiant parabolės aukštį ir plotį, kinta ir morfologiniai filtrai, galima lengviau suprasti filtrų veikimą bei siekiamą rezultatą. Tokia analizė yra tinkama stebėti vienam signalo kanalui minutės ar keletos minučių intervale. Nustačius koeficientus trumpesnei atkarpai, visiems signalo kanalams nuspręsta pasinaudoti rezultatų išvedimu į CSV failą, su pikų skaičiumi kiekvienoje minutėje ir detektavimo riba.

3.2.1. Triukšmo šalinimas signaluose

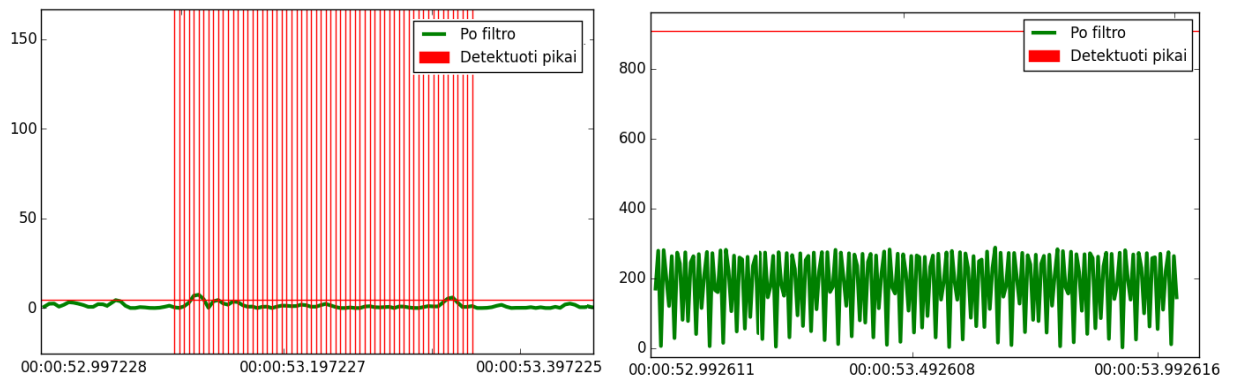
Analizuojant turimą signalų bazę, buvo pastebėta, jog signalai skiriasi savo kokybe. Kai kuriuose signaluose juos vizualiai stebint buvo galima pamatyti tam tikrą triukšmą, kuris ypač trukdo, neturint patirties EEG signalų analizėje. Tiek morfologinių filtrų, tiek gabalais tiesinės funkcijos algoritmų pritaikyme, juose naudojamų koeficientų radimui, dažnu atveju, signalai buvo tiriami vizualiai, todėl signalų nutriukšminimas buvo būtinas.

EEG signaluose gana dažnai galima sutikti įvairių pašalinių veiksnių. Tai paciento mirkčiojimas, staigus sujudėjimas, prakaitavimas, kvėpavimas. Pastarasis signalui galo įnešti žemo dažnio (0.5 Hz) triukšmą [12]. Senesnio modelio EEG prietaisuose žymus yra elektros srovės šaltinio veiksnys, kuris encefalogramas gali užtriukšminti 50 Hz dažnio triukšmu. Šiems triukšmams pašalinti buvo naudojamas Furjė filtras, kuris apdoroja ne pačias signalo reikšmes, o jo spektrą [9]. Taip buvo pašalintos nepageidaujamo dažnio harmonikos iš tiriamų signalų, taip pagerinant signalo kokybę tolimesniems tyrimams. Furjė filtrai buvo įgyvendinti naudojantis Python programavimo kalbos Scipy biblioteka.



9 pav. Kairėje vaizduojamas originalus signalas. Dešinėje signalas, nufiltruotas Furjė filtru. Signalai matuojami mikrovoltais per milisekundę.

Paveikslėlyje matomame viršuje, vaizduojami rezultatai gauti 1 minutę signalo nufiltravus Furjė filtru. Furjė filtras buvo pritaikytas pašalinti 50 Hz ir 0.5 Hz triukšmus. Ne visose tiriamose pacientų encefalogramose yra matomas toks ryškus signalo pasikeitimas po filtrų panaudojimo. Nėra tikslių duomenų ar visos encefalogramos buvo atliktos tuo pačiu prietaisu. Tai ir galėjo lemti skirtingą duomenų kokybę. Tačiau, apdorojus signalus filtru, yra lengviau vizualiai pamatyti epileptoforminius pikus ir taip greičiau nustatyti tinkamas algoritmuose naudojamų parametrų vertes. Morfologinių filtrų algoritmo atveju, didelis triukšmas signale, lemia detekcijos ribos parinkimą. Paveikslėlyje žemiau (10 pav.), galima pamatyti, jog taikant standartinius formulėmis aprašytus struktūrinio elemento parametrus, po morfologinių filtrų panaudojimo (2.8 formulė), signale išlieka nenufiltruotas 50 Hz triukšmas. Šio triukšmo vertės patenka į ekstremumų vidurkių skaičiavimą po nufiltravimo, taip iškreipdamas detekcijos ribą. Statistiškai vidurinė vertė tampa pakankamai aukšta, o kadangi detekcijos riba yra dar bent keletą kartų aukštesnė (2.14 formulė), pikai lieka nedetektuoti. Pateiktame paveikslėlyje, Signalu filtruoto su Furjė filtru detekcijos riba gaunama 4.6, o kairėje vaizduojamame paveikslėlyje, dėl aukšto vidurkio - apie 907 mikrovoltus.



10 pav. Kairėje vaizduojama pikų detekcija prieš tai nufiltravus signalą Furjė filtru. Dešinėje signalas, be Furjė filtro panaudojimo. Signalai matuojami mikrovoltais per milisekundę.

4. Morfologinių filtrų algoritmo įgyvendinimas

Pirmasis įgyvendintas algoritmas remiasi morfologinių filtrų teorija. Šių filtrų pagalba, signalė pašalinama foninė smegenų veikla ir išryškunami epileptoforminiai pikai. Vienas iš šio baigiamojo darbo tikslų - epileptofotminių pikų detekcija, todėl darbas vyko tokiais etapais:

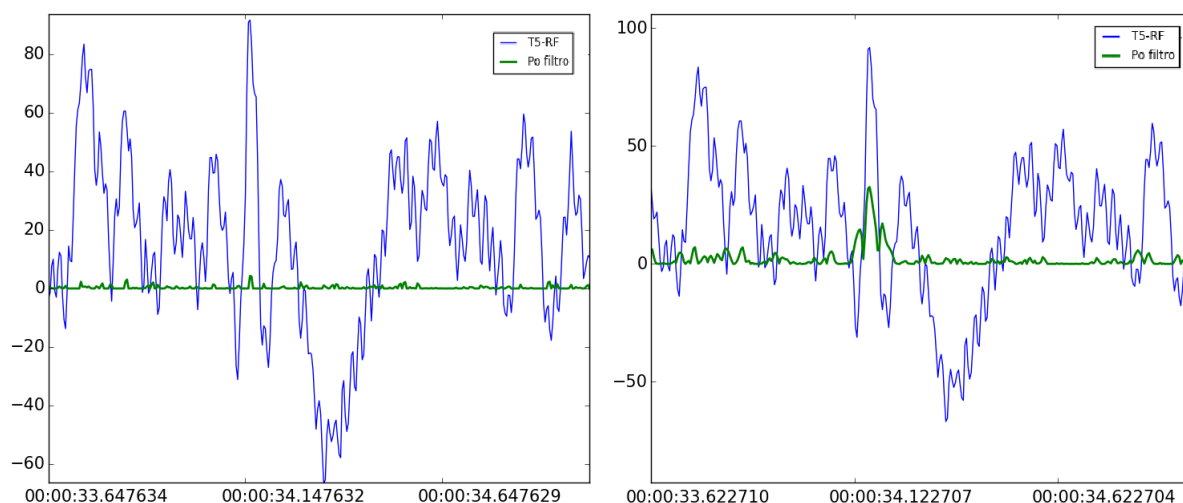
- Encefalogramų nusiskaitymas ir reikiamų kanalų analizei parinkimas.
- Signale esančių smegenų veiklą apibūdinančių parabolų ilgių bei pločių radimas.
- Koeficientų, skirtų sukonstruoti foninę smegenų veiklą apibūdinančią parabolę, apskaičiavimas ir tikslinimas.
- Pikų detekcijos ribos tikslinimas.

Kiekvienai iš turimų encefalogramų, minėtieji etapai buvo atliekami atskirai, taip randant tik tam pacientui būdingus skaičiavimų koeficientus ir detekcijos ribas. Vėliau buvo atlikta analizė, bandant atrasti panašumus ir galimybes pritaikyti standartizuotus parametrus, tinkančius visiems pacientams.

4.1. Struktūrinio elemento ilgio parinkimas

2.2 skyrelio poskyryje buvo aprašyta struktūrinio elemento apskaičiavimo teorija, todėl yra žinoma, jog struktūrinio elemento ilgis, tai parabolę sudarančių taškų skaičius. Kuo struktūrinį elementą sudarys daugiau reikšmių, tuo parabolė bus platesnė, todėl optimalus šio dydžio parinkimas yra būtina sąlyga algoritmo įgyvendinime.

Su testiniais duomenimis buvo pateiktas detalus aprašymas, kurioje signalo minutėje, kiek pikų galima aptikti. Todėl keičiant struktūrinio elemento ilgio reikšmes, buvo tikrinama, ar matomi pikai, atitinka gydytojų pateiktus rezultatus.



11 pav. Kairėje vaizduojamas originalus signalas ir nufiltruotas signalas, kai struktūrinio elemento ilgis yra $0.5 * median|W|$. Dešinėje originalus signalas ir nufiltruotas signalas, kai struktūrinio elemento ilgis yra $6 * median|W|$. Signalai yra matuojami mikrovoltais per milisekundę.

Analizės su realiais pacientų duomenimis metu, buvo pastebėta, jog struktūrinio elemento ilgis turi didelę įtaką, signalo filtravimui. Atlikus analizę ir išbandžius įvairias koeficiento reikšmes, buvo gautas rezultatas, kuris geriausiai išryškina pikus, bei nufiltruoja foninę veiklą:

$$S_L = 4 \text{median}|W|, \quad (4.1)$$

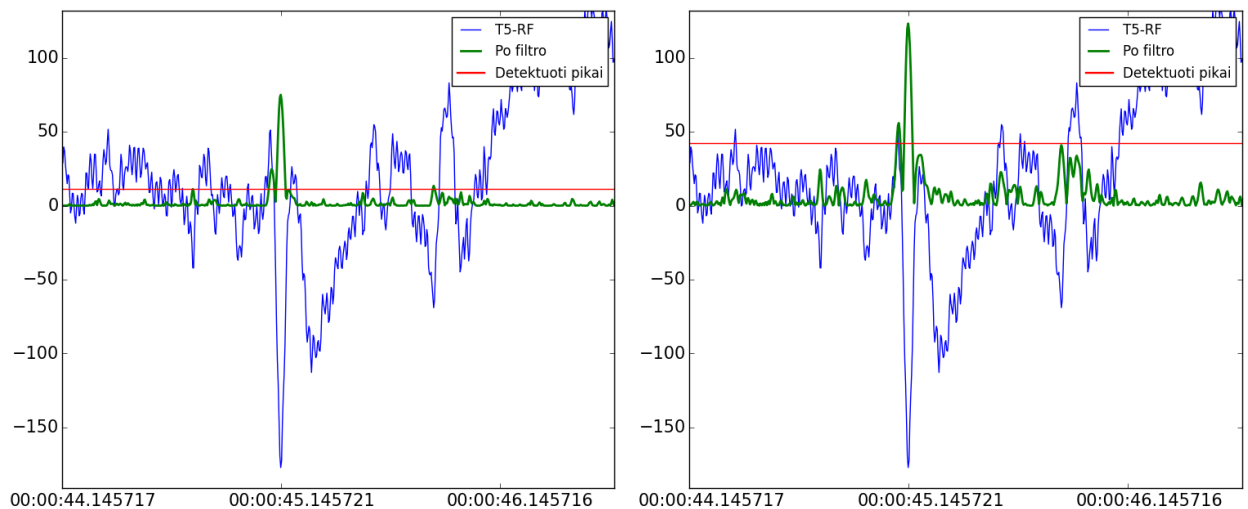
kur S_L - Struktūrinio elemento ilgis, $|W|$ - signalo arkų pločiai.

4.1 formulė apibrėžia, jog struktūrinė parabolė turėtų būti 4 kartus platesnė, nei statistiškai vidurinis encefalogramą sudarančių smegenų veiklos parbolių plotis. Analizuojant testinius duomenis buvo pastebėta, jog dažniausiai smegenų veiklą, aprašančių parbolių pločių mediana svyruoja tarp 9 - 18 taškų, tai atitiktų 0.035 – 0.07 s. Epileptoforminio piko ilgis 0.04 – 0.2s, vadinasi, siekiama užtikrinti, jog struktūrinė parabolė būtų piko pločio.

11 paveikslėlis vaizduoja, jog struktūrinį elementą padarius per trumpu (kairėje), apskaičiuota parabolė yra labai siaura, todėl filtravimo metu, gali būti nufiltruojami ne tik įvairūs triukšmai, bet ir epileptoforminiai pikai. Paveikslėlio dešinėje pusėje demonstruojamas gautas atvirkščias variantas: per ilgas struktūrinis elementas, palieka daug nenufiltruotų įvairių reikšmių.

Reikia paminėti, jog analizės metu, keičiant W reikšmes, kiti struktūrinio elemento parametrai buvo laikomi nekintančiomis konstantomis.

4.2. Struktūrinio elemento aukščio parinkimas



12 pav. Dešinėje vaizduojamas originalus signalas ir signalas po morfologinių filtrų, kai smegenų veiklos parbolių aukščių mediana buvo sumažinta 2 kartus ($|f|/2$). Kairėje vaizduojamas originalus signalas ir signalas po morfologinių filtrų, struktūrinio elemento apskaičiavimui naudojant smegenų veiklos parbolių aukščių medianą. Signalai matuojami mikrovoltais per milisekundę.

Igyvendintame algoritme naudojamos elementariųjų morfologinių operacijų - morfologinio išplėtimo ir susiaurinimo - įvairios kompozicijos. Jos vykdomos su susidarytu struktūriniu elementu - parabole, kurios koeficientai randami apskaičiuojant vidutines visų signalo esančių parbolių aukščių bei pločių reikšmes. Galutinis filtras (2.7 formulė), tai vidurkis paėiliui vykdomų morfologinio atidarymo ir uždarymo, bei uždarymo ir atidarymo. Norint patikslinti ir suprasti algoritme

naudojamų parametru reikšmes, pirmiausia svarbu išsiaiškinti kaip rasti foninės smegenų veiklos parbolių aukščiai bei pločiai veikia pačią struktūrinę parabolę.

Struktūrinė parabolė gaunama koeficientų a ir b dėka. b - tai foninės smegenų operacijas aprašančių parbolių aukščių mediana, kuri tiesiogiai savo reikšme didina visą struktūrinę parabolę. a - tai santykis tarp parbolių aukščių medianos ir parbolių pločių medianos. 2.10 ir 2.11 formulėse matoma, kad šių koeficientų apskaičiavimui didelę įtaką turi foninės smegenų veiklos parbolių aukštis $|f|$. Bandymais buvo pastebėta, jog mažinant foninių parbolių aukštį, žemėja pati struktūrinė parabolė, todėl buvo nuspręsta atlikti detalesnį tyrimą $|f|$ nustatymui.

Praėjusio semestro metu, moksliniame tiriamajame darbe, buvo nustatyta, jog optimaliai struktūrinė parabolė yra sukonstruojama, kai foninės smegenų veiklos parbolių aukštis $|f|$ yra dalijamas pusiau. Taip gaunama žemesnė struktūrinė parabolė, kuri labiau išryškina epileptoforminius pikus. Tačiau atlikus detalesnį tyrimą baigiamajame magistro darbe, nuspręsta parbolių aukščio $|f|$ nemažinti, nes taip padidinama nufiltruoto signalo amplitudė ir tiesiogiai veikiama epileptoforminių pikų detekcijos riba. 12 paveikslėlyje vaizduojamas originalus signalas, kurio filtravimui buvo panaudota parabolė, kurios aukščio skaičiavimui naudota matomo signalo atkarpos parbolių mediana - 14 mikrovoltų (dešinėje) ir pusė medianos - 7 mikrovoltai (kairėje). Pirmuoju atveju, randami du pikai, antruoju, labiau išryškinami pikai, tačiau, norint detektuoti antrąjį piką, tektų keisti detekcijos ribos parametą.

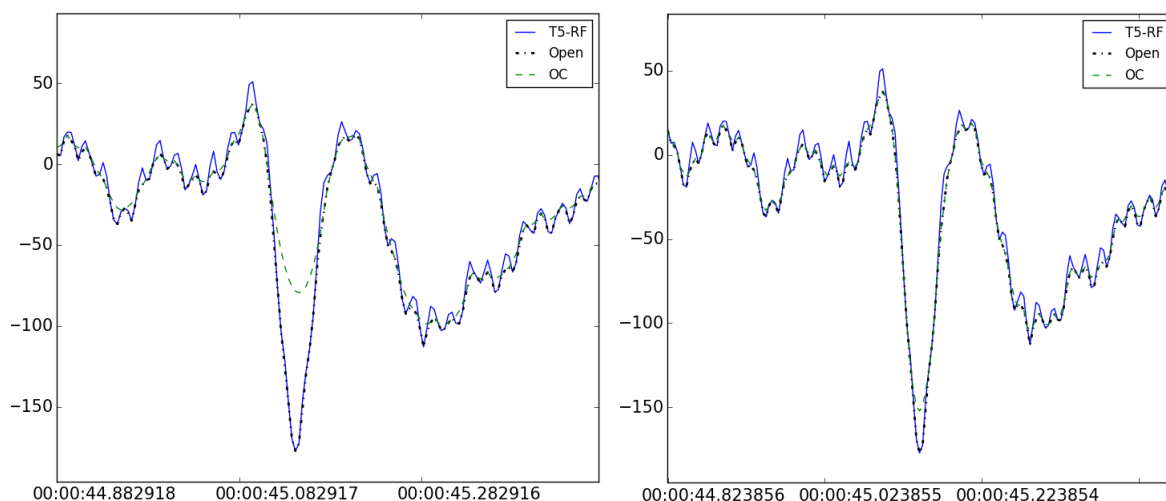
4.3. Struktūrinių parbolių skirtumas

2.9 formule aprašomos dvi parabolės g_1 ir g_2 , kurios yra struktūriniai elementai morfologinėse operacijose. CO ir OC filtruose (2.5 ir 2.6 formulės) paeilui yra atliekamos morfologinio uždarymo ir atidarymo operacijos su abiem struktūriniais elementais. Kaip jau žinoma, struktūrinės parabolės g_1 ir g_2 yra tokio paties pločio, tačiau skiriasi jų aukščiai. Aukščių skirtumą lemia koeficiento a radimo formulės:

$$a_1 = \frac{\text{median}|f|}{0.5\text{median}|W|}, \quad a_2 = \frac{\text{median}|f|}{1.5\text{median}|W|} \quad (4.2)$$

Koeficientas a_1 yra 3 kartus didesnis už koeficientą a_2 . Tai lemia ir struktūrinių elementų g_1 ir g_2 aukščių skirtumą. Šis skirtumas yra parenkamas ne veltui. Vykdamt morfologines operacijas su skirtingų aukščių struktūrinėmis parabolėmis, labiau išryškinami epileptoforminiai pikai (13 pav.).

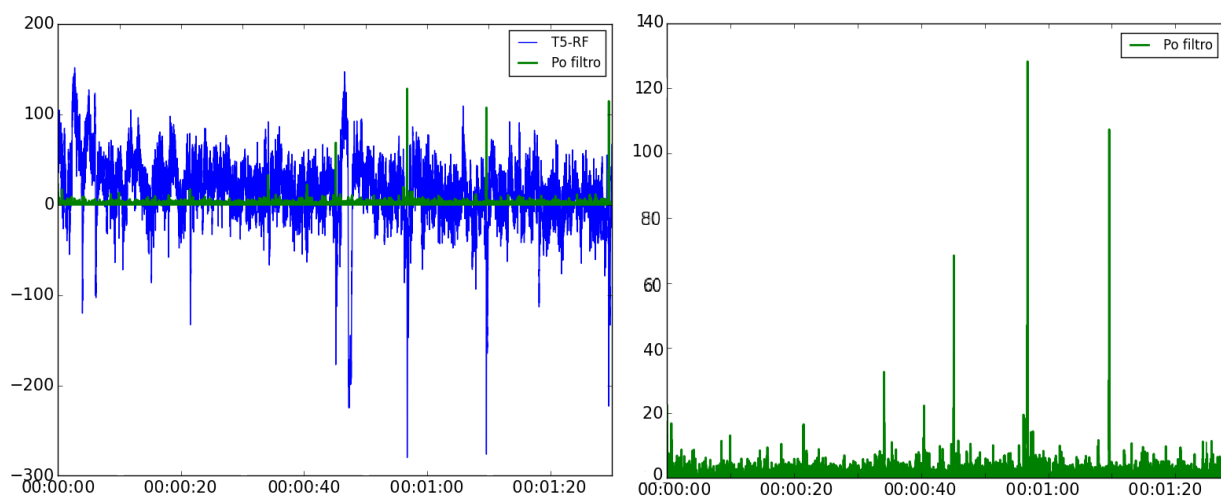
13 paveikslėlis vaizduoja, jog po OC filtro, įvykdyto su vienodo aukščio struktūriniais elementais (dešinėje), epileptoforminis pikas yra labiau išpildomas gautuoju signalu. Morfologinis atidarymas yra vykdomas su pirmuoju struktūriniu elementu g_1 . Matoma, jog po filtravimo, gautasis signalas yra vizualiai panašus į piką. Po to vykdomas morfologinis uždarymas su struktūriniu elementu g_2 . Kadangi antrasis struktūrinis elementas yra vienodo aukščio, po filtravimo skirtumas tarp originalaus signalo ir gautojo yra pakankamai nežymus. Kadangi galutinis filtras yra OC ir CO filtrų sumos vidurkis, gaunamas rezultatas būtų labai panašus savo aukščiu į piką. Atėmus OC-CO filtrą iš realaus signalo, gautasis skirtumas būtų labai mažas ir epileptoforminis pikas nebūtų toks ryškus detekcijai. Paveiklėlio kairėje vaizduojamas filtravimas su skirtingų aukščių struktūrinėmis parabolėmis, todėl po OC filtro, išryškėja skirtumas tarp signalų ir taip garantuojamas didesnis patikimumas piko radimui.



13 pav. Kairėje vaizduojamas originalus signalas ir nufiltruoti signalai OC bei morfologinio atidarymo filtrais, kai struktūrinio elemento parametrai a_1 ir a_2 skiriasi 3 kartus. Dešinėje originalus signalas ir nufiltruoti signalai OC bei morfologinio atidarymo filtrais, kai struktūrinio elemento parametrai a_1 ir a_2 yra vienodi. Signalai yra matuojami mikrovoltais per milisekundę.

4.4. Pikų detekcijos ribos skaičiavimas

Įvykdžius morfologinių filtrų operacijas, apskaičiuojamas galutinis signalas - tai skirtumas tarp originalaus signalo prieš morfologinių filtrų panaudojimą ir OCCO filtru apdoroto signalo (2.8 formulė). Gautasis signalas yra mažesnės amplitudės nei originalusis ir jame yra išryškinti pikai.



14 pav. Kairėje vaizduojamas originalus signalas ir originalaus bei nufiltruoto OCCO filtru skirtumas. Dešinėje išdidintas originalaus bei nufiltruoto OCCO filtru signalų skirtumas. Signalai yra matuojami mikrovoltais per milisekundę.

Paveikslėlio viršuje dešinėje, matomas nufiltruoto signalo modulis, taip apsiribojant tik teigiamomis reikšmėmis. Akivaizdu, jog vienos signalo viršūnės yra didesnės už kitas, tačiau riba, nuo kurios aukščiau jos yra laikomos epileptoforminiu piku neaiški.

Norint detektuoti signale pikus, reikia pasinaudoti 2.14 formule ir apskaičiuoti, tiesę, nuo kurios visos ją viršijusios signalo viršūnės, bus laikomos ligos ženklų. Šios ribos nustatymui reikia rasti nufiltruotojo signalo viršūnių amplitudžių vidurkį ir padauginti iš detekcijos ribos koeficiento d . Detekcijos ribos koeficiento d radimas yra vienas iš šio magistro baigiamojo darbo tikslų, todėl jam buvo atliktas išsamesnis tyrimas.

Kiekvienai iš testinių encefalogramų buvo atliekamas morfologinis filtravimas ir gaunamas galutinis signalas (2.8 formulė). Kadangi gydytojai buvo pateikę konkretų epileptoforminių pikų skaičių, kiekvienoje iš 20 signalo minučių, rankiniu būdu, buvo keičiamas algoritme ribos koeficientas d . Taip randama konkreti koeficiento reikšmė, būdingas tik tuo metu tiriamam pacientui.

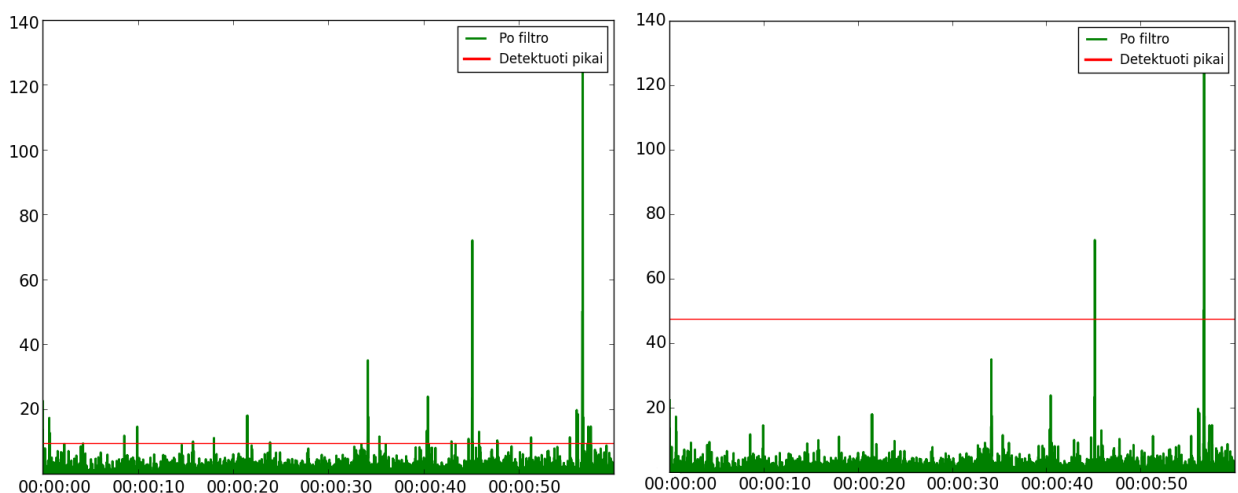
4.4.1. Encefalogramų tyrimui parinkimas

Pikų detekcijos ribos koeficiento d tikslinimo metu, buvo pastebėta, jog šis skaičius yra skirtingas kiekvienai iš tiriamų encefalogramų. Rastasis koeficientas svyravo tarp 5.5 – 15. Dėl šitokio neapibrėžtumo, nuspręsta, peržiūrėti encefalogramas, kurios turėjo minimalią ir maksimalią reikšmes.

Itin žemą detekcijos ribos koeficientą turėjusi encefalograma pasižymėjo itin dideliu pikų skaičiumi ir joje akivaizdžiai buvo ryškus 50 Hz triukšmas (9 pav. kairėje). Gydytojas atliko šios encefalogramos tyrimą be panaudoto Furjė filtro, todėl buvo padaryta prielaida, jog detektuojamų pikų skaičius, ir gydytojų nustatytas skaičius, gali skirtis. Dėl šios priežasties ši encefalograma buvo pašalinta iš tolimesnio ribos koeficiento tyrimo.

Maksimalias ribos detekcijos koeficiento reikšmes turėjusios keletas encefalogramų pasižymėjo itin mažu gydytojų detektuotų pikų skaičiumi (mažiau nei 10 pikų visame signale). Todėl šių keletos pikų radimui, ribos koeficientas turėdavo būti itin aukštas ir taip labai išsiskirdavo iš likusių rezultatų.

4.4.2. Detekcijos ribos koeficiento tikslinimas



15 pav. Vaizduojamas nufiltruotas signalas ir detekcijos riba, kai detekcijos ribos koeficientas d yra 8 (kairėje) ir 16 (dešinėje). Signalai yra matuojami mikrovoltais per milisekundę.

Tikslus ribos detekcijos koeficiento nustatymas yra svarbus žingsnis encefalogramų tyrime. Jei koeficientas bus per mažas, po morfologinių filtrų algoritmo veikimo, bus detektuojamas per

didelis epileptoforminių pikų skaičius. Jei per didelis, ne visi pikai bus aptinkami (15 pav.).

Testavimo metu nufiltravus visus tiriamų pacientų signalus ir pašalinus encefalogramas, kuriose matomas triukšmas bei encefalogramas su labai mažu epileptoforminių pikų skaičiumi, buvo paskaičiuotas vidurkis tarp visų gautųjų detekcijos ribos koeficientų ir gauta reikšmė:

$$d^* = 9.7 \quad (4.3)$$

Gautasis skaičius yra vidurkis visų prieš tai rastųjų koeficientų, todėl buvo paskaičiuota, kaip šis ribos detekcijos koeficientas keičia rezultatus, pritaikius jį filtravimui visiems tiriamiems signalams.

Signal.	Gyd. detekt. pikai	$d^* - 20\%$	$d^* - 10\%$	d^*	$d^* + 10\%$	$d^* + 20\%$	$d^* + 25\%$
1	166	74.09	44.57	13.85	-5.42	-24.69	-30.72
2	132	65.15	31.81	13.63	-1.51	-11.36	-14.39
3	163	169.93	119.63	71.77	41.71	6.13	-3.68
4	158	163.92	100.63	65.18	32.27	8.22	0
5	207	142.99	75.84	29.46	-9.66	-33.81	-43.96
6	857	-3.15	-20.77	-34.18	-45.27	-57.99	-62.42
7	370	-1.89	-31.35	-51.62	-63.51	-72.97	-76.75
8	152	43.42	2.63	-22.89	-50	-61.84	-64.47
9	398	52.01	22.11	-3.51	-24.12	-38.19	-44.72

3 lentelė. 9 pacientų encefalogramų duomenys procentais, nufiltravus jas morfologiniais filtrais, kai detekcijos ribos koeficientas d^* yra didinamas 10%, 20%, 25 % ir mažinamas 10%, 20%. Taip pat matomi iš anksto gydytojų pateikti epileptoforminių pikų skaičiai tiriamame kanale.

Lentelėje pateiktoje viršuje, galima matyti, kaip keičiasi skirtingų pacientų encefalogramų tyrimo duomenys, parenkant vis kitą detekcijos ribos koeficientą d . Rezultatai pateikiami, laikant jog gydytojų nustatytas pikų skaičius yra 100 %, taip randamas procentinis skirtumas tarp šio skaičiaus ir algoritmu rastų pikų skaičiaus. Teigiami rezultatai reiškia, kiek procentaliai buvo rasta pikų virš gydytojų nustatytos ribos. Neigiami - koks procentinis trūkumas iki gydytojų ribos. Tamsiau paryškinti langeliai, kuriuose rezultatas yra artimas tikrajam pikų skaičiui encefalogramoje. Kai ribos koeficientas d prilyginamas bandymų būdu rastam vidurkiui d^* , tik vienam iš tiriamųjų signalų, detektuojamas panašus į gydytojų pateiktą epileptoforminių pikų skaičių, rezultatas - apie 3.5 % trūkumas (384 pikai).

Todėl galima daryti išvadą, jog detekcijos ribos koeficientas turėtų svyruoti aplink savo vidurkinę reikšmę d^* . Taip būtų sudaromos sąlygos gauti epileptoforminių pikų režius nuo minimalaus iki maksimalaus pikų skaičiaus. Imant rezultatus nuo -20 % iki +25 % nuo d^* , gaunama išvada, jog pikų detekcijos riba turėtų būti tarp 7.76 – 12.12.

5. Interpoliavimo gabalais tiesine funkcija algoritmo įgyvendinimas

Antrasis algoritmas, skirtas epileptoforminių pikų detekcijai, remiasi interpoliavimu gabalais tiesine funkcija: per pasirinktą atkarpą yra išvedamos tiesės, artimiausios signalui mažiausių kvadratų metodu prasme. Remiantis tam tikromis gabalais tiesinę funkciją sudarančių tiesių savybėmis, atrenkamos signalo atkarpos ir daroma prielaida, jog tai epileptoforminiai pikai.

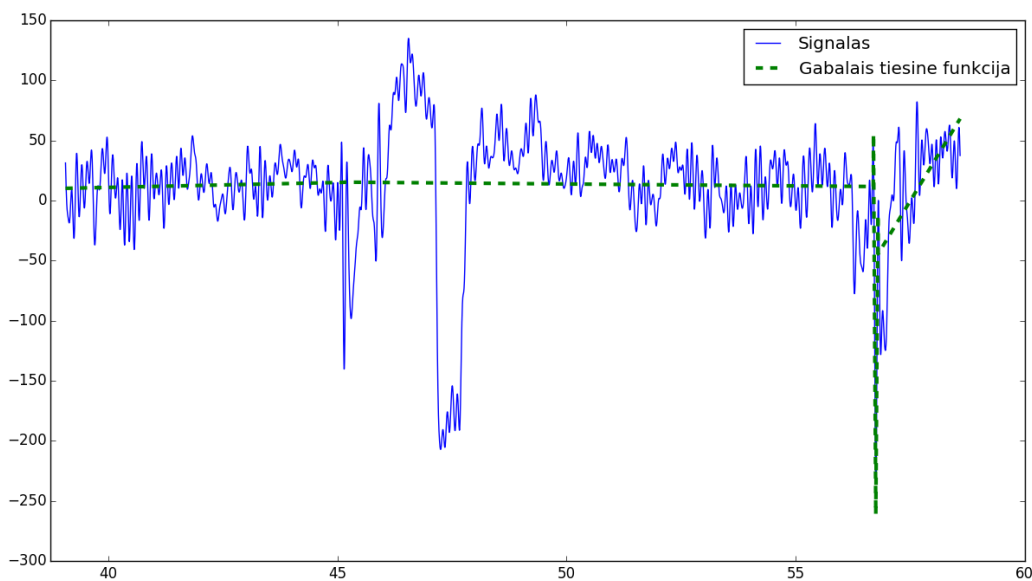
Algoritmas buvo suprogramuotas atvirojo kodo Python programavimo kalba. Realizacijoje daugiausia buvo naudojamos SciPy biblioteka ir joje esančiomis interpoliavimo funkcijomis.

Algoritmo įgyvendinimo etapai:

- Encefalogramų nusiskaitymas ir reikiamų kanalų analizei parinkimas.
- Interpoliavimo gabalais tiesine funkcija vykdymas.
- Gautųjų galimų pikų filtravimas remiantis tiesių krypties koeficientų savybėmis.
- Tikslios piko vietos signale pažymėjimas.

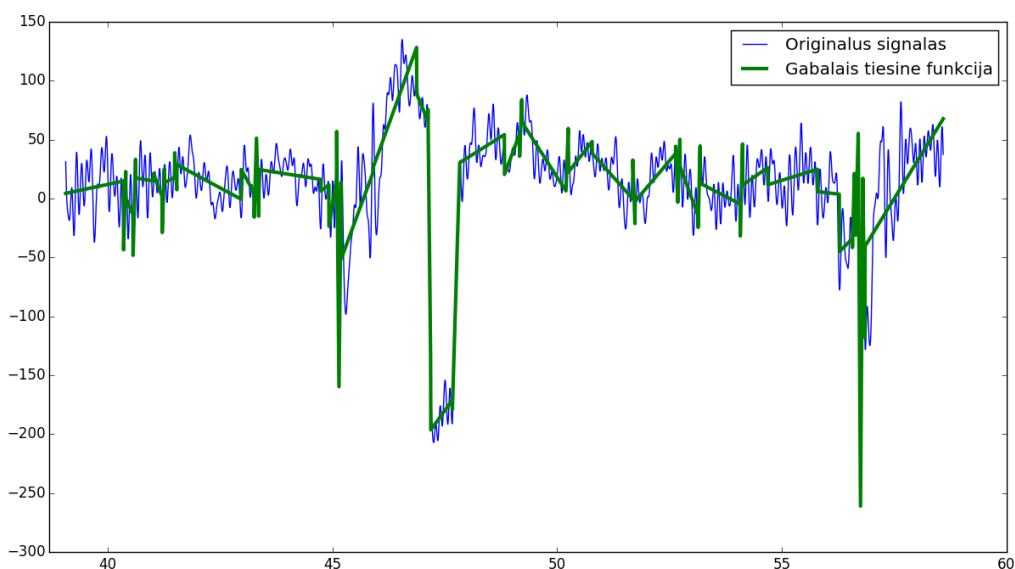
5.1. Intervalo pasirinkimas interpoliavimui

Nagrinėjamas encefalogramas sudaro 21 kanalas, kuris kiekvienas talpina apie 20 minučių paciento smegenų veiklos informacijos. Gabalais tiesinės funkcijos pritaikymas vienam kanalui reikalauja signalą skirstyti trumpesniais intervalais. Jei gabalais tiesine funkcija interpoliuotume visą signalo laiką, būtų gautos 5 tiesės, kurios teoriškai turėtų atvaizduoti vieną epileptoforminį piką. Nors signale įmanomas ir tik vieno epileptoforminio piko buvimas, tačiau žinant, jog piko trukmė tėra 40 - 200 ms [7], bei turimoje duomenų bazėje vidutinis pikų skaičius yra 251 pikas visame signale, akivaizdu, jog signalas turi būti skaidomas mažesniais tiriamais intervalais.



16 pav. Vaizduojamas originalus signalas ir signalas, gautas pritaikius interpoliavimą gabalais tiesine funkcija visam matomam laiko intervalui. Signalai matuojami mikrovoltais per milisekundę.

Ilustracijoje aukščiau vaizduojamas atvejis, kai tiriamas intervalas interpoliavimui gabalais tiesine funkcija yra per didelis. Matomose 20 sekundžių, nubraižomos penkios tiesės, taip tikintis gauti epileptiforminį piką, tačiau mažesni pikai lieka neaptikti (ties 45 sekunde). Skaidant signalą intervalais, kuriuose atliekamas interpoliavimas gabalais tiesine funkcija, galimas net ir mažesnių epileptiforminių pikų aptikimas. 17 paveikslėlyje matoma ta pati 20 sekundžių signalo atkarpa, tačiau gabalais tiesinė funkcija buvo pritaikyta kas 2 sekundes, taip atvaizduojant daugiau signale esančių viršūnių. Gautajame rezultate, vizualiai galime matyti, jog prieš tai nepakliuvę galimi pikai, mažesniuose intervaluose dominuoja ir išlieka tolimesniems tyrimams.

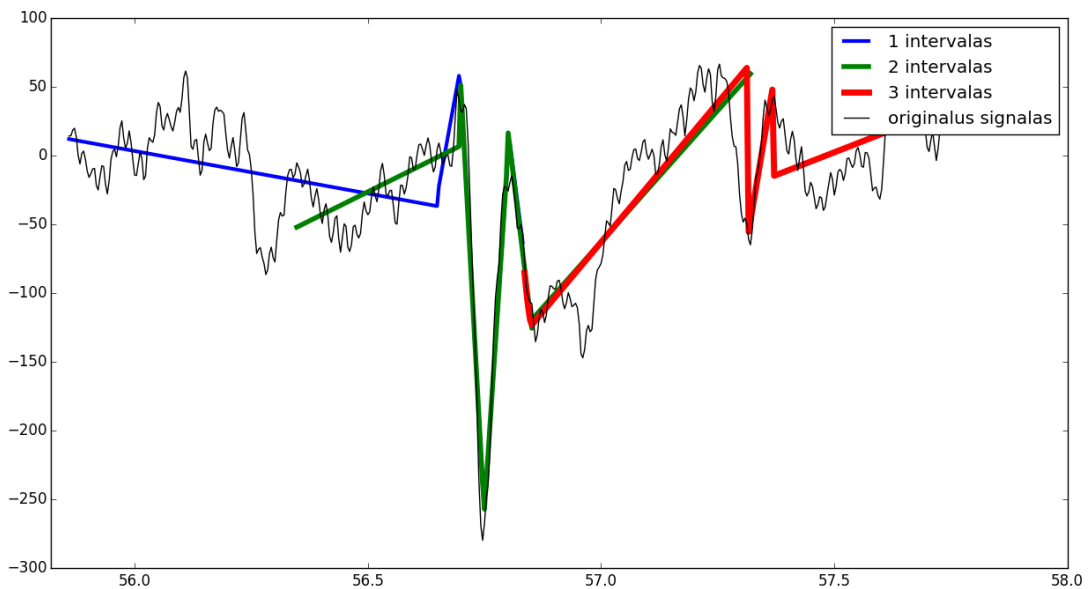


17 pav. Vaizduojamas originalus signalas ir signalas, gautas pritaikius interpoliavimą gabalais tiesine funkcija atkarpomis kas 2 sekundes. Signalai matuojami mikrovoltais per milisekundę.

Po atliktų testavimų su pacientų encefalogramomis, buvo nustatyta, jog optimalus intervalas interpoliacijai yra 400 ms. Kadangi maksimali epileptiforminio piko trukmė yra 200 ms, intervalas, kuriame bus vykdomas interpoliavimas gabalais tiesine funkcija, turi būti ilgesnis, norint jog pirmoji ir penktoji tiesės iš 5 tiesių rinkinio, rodytų normalią smegenų veiklą.

5.2. Žingsnio pasirinkimas interpoliavimui

Skaidant turimą signalą intervalais, išryškėjo problema, jog pasirinkus atkarpą, kurioje pritaikomos 5 tiesės, ne tokie ryškūs pikai bus neįtraukiami. Kiekviename intervale, kuriame vykdomas interpoliavimas gabalais tiesine funkcija, nubraižoma vienas tiesių rinkinys, atitinkantis tame intervale ryškiausią piką. Jei antro epileptiforminio piko pradžia ar pabaiga yra tiriamame intervale, kitame intervale, pikas galimai nebus aptiktas. Todėl signalą suskaidę intervalai turi persidengti. Programoje buvo pasirinktas intervalo ilgis, kuriame atliekama interpoliacija ir žingsnis, kuriuo šis intervalas yra slenkamas. Atlikus interpoliavimą gabalais tiesine funkcija pirmame intervale, žingsnis yra pridamas prie pirmojo intervalo pradžios, taip gaunant antrą naują interpoliavimui skirtą intervalą. Dalis signalo taškų priklausomai nuo žingsnio pasirinkimo priklausys net keletui intervalų. Taip yra užtikrinama, jog nei vienas epileptiforminis pikas neliktų nedetektuotas dėl tiesiog blogo intervalų pasirinkimo algoritme.



18 pav. Vaizduojamas originalus signalas ir signalas, gautas pritaikius interpoliavimą gabalais tiesine funkcija intervalais kas sekundę su 0,5 s žingsniu. Signalai matuojami mikrovoltais per milisekundę.

18 paveikslėlis iliustruoja aprašytą minėtąjį procesą: pasirinkus intervalą interpoliacijai, jis yra slenkamas norimu žingsniu per signalą. Šiuo atveju vaizduojamos 2 signalo sekundės, kuriose intervalas interpoliavimui yra 1 sekundė, o slinkimo žingnis 0,5 sekundės. Taip gauname 3 persidengiančius intervalus.

Po atliktų testavimų su pacientų encefalogramomis, buvo nustatyta, jog optimalus žingsnis interpoliacijai 80 ms. Žingsnis buvo parinktas bandymų keliu, bet ne didesnis, nei skaičiuojamas intervalas. Kuo mažesnis žingsnis, tuo gaunami tikslesni rezultatai, tačiau tai didina iteracijų skaičių algoritme, taip apkraudamas ir lėtindamas programos veikimą. Todėl buvo gautas optimalus 20 žingsnių kode rezultatas, kas atitinka 80 signalo milisekundžių.

5.3. Epileptoforminių pikų detekcija

5.3.1. Tiesių koeficientų svarba

16 ir 17 paveikslėliuose galima pastebėti, jog mažinant intervalus interpoliavimo atlikimui, gautasis signalas vis labiau pradeda atitikti originalųjį. Optimaliai parinkus intervalą ir žingsnį, visi potencialūs epileptoforminiai pikai yra atvaizduojami naujai gautame funkcijos grafike, tačiau kartu yra atvaizduojama ir tiesiog neįprasta žmogaus smegenų veikla. Tai sugeneruoja daug pikų, todėl epileptoforminius pikus rasti padeda gabalais tiesinės funkcijos tiesių koeficientai k_1, k_2 .

Kadangi yra žinoma epileptoforminio piko forma (6 paveikslėlis), tai pirmųjų tiesių f_1, f_2 iš gabalais tiesinės funkcijos koeficientai k_1, k_2 turi skirtingus ženklus. Pirmoji tiesė f_1 - tai normali smegenų veikla iki piko. Ši tiesė turi neigiamą koeficientą k_1 . Antroji tiesė f_2 - tai piko pradžia, ji žymi staigų smegenų aktyvėjimą ir tiesės kilimą aukštyn, todėl jos koeficientas k_2 yra teigiamas. Remiantis koeficientų filtrais aprašytais 2.3 skyrelyje, visiems sugeneruotiems pikams yra pritaikomos šios pikų savybės :

- $|k_1| * |k_2| < 0$,

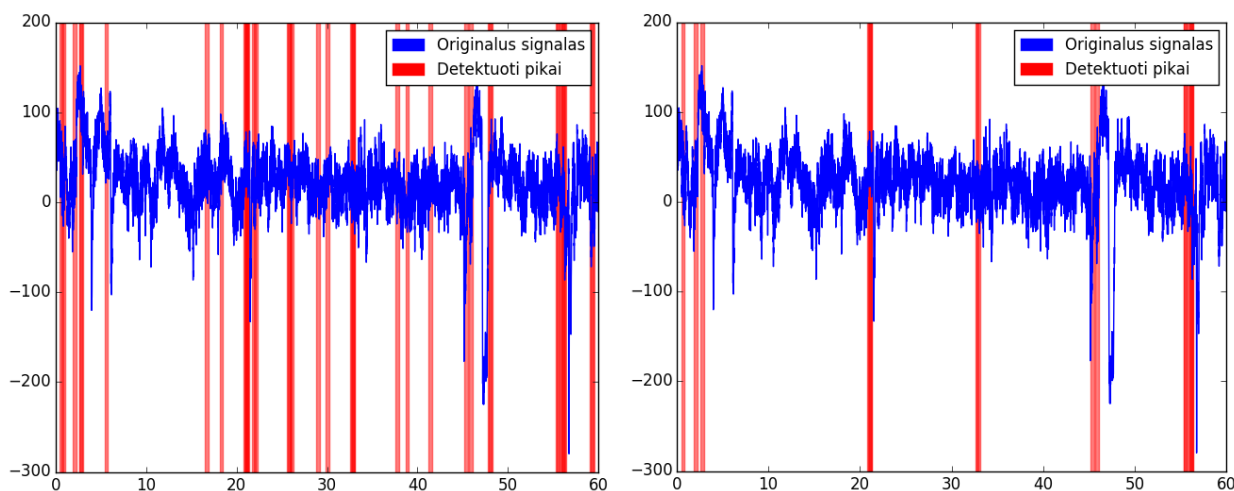
- $||k_1| - |k_2|| > const.$

Pirmoji piko savybė nurodo, jog piku laikomas tik toks tiesių rinkinys, kurių pirmųjų krypties koeficientų kryptys yra skirtingos. Antroji savybė - tai piko kampų lyginimas. Krypties koeficientas k_j , $j = 1, 2$ nurodo kampą, kuriuo kyla nurodytoji tiesė. Epileptoforminiams pikams būdingas staigus kylimas, todėl ir skirtumas gaunamas pakankamai didelis. Algoritmo testavimo metu, pritaikius pirmąją savybę tiriamam signalui, surastų galimų pikų skaičius vidutiniškai sumažėdavo ketvirtadaliu.

5.3.2. Tiesių koeficientų skirtumo parinkimas

Interpoliuojant signalą gabalais tiesine funkcija, gaunamas nemažas netikrų epileptoforminių pikų skaičius. Todėl vienas iš pagrindinių tikslų, įgyvendinant šį algoritmą, yra tiesių koeficientų k_1, k_2 modulių skirtumo nustatymas (2.23 formulė), kuris leistų detektuoti, tik tuos pikus, kuriems būdingas staigus antrosios tiesės kylimas. Taip ieškoma, skaičiaus, kuris apibūdintų skirtumą, tarp dviejų tiesių nuolydžio kampų, taip, jog galėtume teigti, jog visi koeficientų skirtumai didesni nei ieškoma konstanta yra epileptoforminiai pikai.

Iš gydytojų pateiktų pacientų signalų, buvo pasirinktas vienas tiriamas signalas. Iš anksto buvo žinoma, kuriame encefalogramos kanale yra nustatyti pikai ir tikslus jų skaičius visame kanalo laike. Taip buvo ieškoma, su kuria konstanta galima gauti gydytojų norodytą pikų skaičių. Kadangi algoritme imamos tik absoliučios koeficientų k_1, k_2 reikšmės, šių koeficientų modulių skirtumas gaunamas gana didelis teigiamas skaičius. Nustačius ribą vienam pacientui, su kuria visoje encefalogramoje gaunamas artimas skaičius pateiktajam iš anksto žinomam gydytojų nustatytam pikų skaičiui, buvo tiriama kito paciento encefalograma, neatsižvelgiant į prieš tai gautuosius rezultatus.



19 pav. Vaizduojamas originalus signalas su detektuotais pikais, kai tiesių koeficientų skirtumas $||k_1| - |k_2|| > 1700$ (dešinėje) ir $||k_1| - |k_2|| > 2300$. Signalai matuojami mikrovoltais per milisekundę.

Paveikslėlyje viršuje, matoma, jog tiesių koeficientų k_1 ir k_2 modulių skirtumo modulis turi būti nustatytas gana tikslus. Parinkus blogą konstantą, kai kurie epileptoforminiai pikai gali likti neaptikti. Tyrimo metu paaiškėjo, jog pacientams interpoliavimo gabalais tiesine funkcija algoritme naudojamų tiesių koeficientų modulių skirtumų modulis yra labai įvairus ir jo reikšmės gali

svyruoti priklausomai nuo signalo. Gautos reikšmės svyravo nuo 650 iki 3000, taigi skyrėsi 4.6 karto. Išsivedus kelių signalų $||k_1| - |k_2||$ skirtumo rezultatus į CSV failą ir stebint reikšmes rankiniu būdu, buvo pamatyta, jog šio skirtumo mediana yra gana pastovus skaičius kiekvienam signalui ir ieškoma konstanta gali būti randama panaudojus $||k_1| - |k_2||$ medianą :

$$c = (\text{median}(|k_1| - |k_2|)) * m, \quad (5.1)$$

kur $||k_1| - |k_2|| > c$ ir $m > 0$

Ieškomas koeficientas m nurodo, kiek kartų tiesių koeficientų k_1 , k_2 skirtumo modulis yra didesnis už savo medianą. Testavimo su realių pacientų encefalogramomis metu, buvo gauti m rezultatai, kurie svyravo nuo 1.3 iki 4.5. Nors koeficiento m įvedimas palengvino tiesių koeficiento c radimą, bei sumažino santykį tarp gaunamų reikšmių iki 3.5 karto, bet vis vien sukėlė didelį neapibrėžtumą. Todėl nuspręsta išsamiau ištirti, ar visi signalai gali būti naudojami tolimesniems skirtumo koeficientų vertinimams.

5.3.3. Encefalogramų tyrimui parinkimas

Didžiajai daliai encefalogramų, turimų testinėje duomenų bazėje, pritaikius interpoliavimo galais tiesine funkcija algoritmą, tiesių koeficientų k_1 ir k_2 modulių skirtumo modulis gaunamas gana panašus skaičius, išskyrus keletą išsiskiriančių encefalogramų. Kaip ir morfologinių filtrų algoritmo atveju, maksimali konstanta buvo gauta signalams, kuriuose gydytojai detektavo vos keletą pikų. Minimali konstanta signalams, kuriuose detektuotų pikų skaičius apie 900-1000 pikų. Viename iš šių signalų, gana ryškus triukšmas (9 pav. kairėje), tačiau kito signalo kokybė nebuvo tokia pastebimai bloga. Taip buvo pastebėtas vienas šio algoritmo minusų, jeigu epileptoforminių pikų skaičius yra labai didelis, epileptoforminiai pikai yra mažesnės amplitudės ir tiesių koeficientų k_1 ir k_2 modulių skirtumo modulis net ir pikų atveju tampa panašus į foninę smegenų veiklą. Šiuo konkrečiu tiriamu atveju, konstanta m buvo gauta 1.3 ir tai reiškė, jog k_1 ir k_2 skirtumas tik 1.3 karto turėtų būti didesnis už vidutinę gaunamą k_1 ir k_2 skirtumo reikšmę, jog visi didesni rezultatai būtų laikomi epileptoforminiais pikais.

Šio netikslumo pataisymui buvo bandyta medianą keisti vidurkiu, tačiau tai nedavė geresnių rezultatų. Nustačius, jog signalai, kuriuose yra itin mažas ir itin didelis pikų skaičius gali iškreipti koeficientus, buvo nuspręsta jų neįtraukti į tolimesnius skaičiavimus.

5.3.4. Koeficiento m tikslinimas

Testavimo metu nufiltravus visus tiriamų pacientų signalus ir pašalinus encefalogramas, kuriuose matomas triukšmas, labai didelis pikų skaičius bei encefalogramas su labai mažu epileptoforminių pikų skaičiumi, buvo paskaičiuotas vidurkis tarp visų m koeficientų ir gauta reikšmė:

$$m^* = 2.9 \quad (5.2)$$

Gautasis skaičius yra vidurkis visų prieš tai rastųjų koeficientų, todėl buvo paskaičiuota, kaip šis daugiklis keičia rezultatus, pritaikius jį filtravimui visiems tiriamiesiems signalams.

Signal.	Gyd. detekt. pikai	$m^* - 25\%$	$m^* - 15\%$	$m^* - 5\%$	$m\%$	$m^* + 5\%$	$m^* + 15\%$
1	166	92.77	53.01	24.69	15.66	4.21	-9.03
2	132	164.39	97.72	40.15	26.51	-1.51	-30.3
3	163	103.68	66.87	23.31	16.56	-3.06	-20.85
4	158	149.3671	103.16	41.77	27.84	5.69	-19.62
5	207	106.28	58.45	14	-0.48	-16.9	-14.97
6	152	131.57	92.76	54.6	44.07	36.84	-0.65
7	398	-2.51	-26.88	-43.96	-49.74	-54.52	-66.33
8	370	-1.08	-20	-41.08	-46.21	-55.4	-66.21

4 lentelė. 8 pacientų encefalogramų duomenys procentais, nufiltravus jas interpoliavimo gabalais tiesine funkcija algoritmu, kai tiesių statumo koeficientas m^* yra mažinamas 5%, 15%, 25 % ir didinamas 5%, 15%. Taip pat matomi iš anksto gydytojų pateikti epileptoforninių pikų skaičiai tiriamame kanale.

Pateiktoje lentelėje viršuje matomi encefalogramų duomenys, nufiltravus jas interpoliavimo gabalais tiesine funkcija algoritmu, kai koeficientas m yra didinamas arba mažinamas. Pateikiami iš anksto gydytojų nustatytas pikų skaičius, kuris yra laikomas 100 %. Taip randamas procentinis skirtumas tarp šio skaičiaus ir algoritmu detektuotų pikų skaičiaus. Neigiamos reikšmės - tai nedektuotų pikų skaičius procentais. Teigiami procentai, tai per daug rastų pikų procentinė išraiška. Kai koeficientas m prilyginamas prieš tai rastųjų koeficientų vidurkiui m^* , vienam ketvirtam pacientui randamas epileptoforninių pikų skaičius labai panašus į gydytojų detektuotąjį. Trūkumas -0.48 - tai 206 rasti pikai. Paryškinti langeliai nurodo, rastą pikų skaičių, artimiausiam gydytojų nustatytam skaičiui.

Matoma lentelė iliustruoja, jog vieno tinkamo visiems pacientams tiesių skirtumo koeficiento rasti nepavyko. Tiriamu atveju, jis svyruoja nuo 2.2 iki 3.3.

6. Algoritmų palyginimas ir rezultatai

Šiame darbe buvo įgyvendinti du algoritmai epileptoforminių pikų detekcijai. Pirmasis remiasi morfologinių filtrų teorija, antrasis interpoliavimu gabalais tiesine funkcija. Algoritmų testavimo metu, abiejų minėtųjų algoritmų veikimas ir patikimumas buvo bandomi tokiais pačiomis sąlygomis: naudojant tą patį kompiuterį² ir atliekant analizę su tokiais pačiomis encefalogramomis. Vertinti algoritmų vykdymą laiko prasme, yra labai sudėtinga. Nors atliekami žingsniai gana panašūs: abu algoritmai vykdydavo duomenų nusiskaitymą iš encefalogramų, filtravimą Furjė filtru, filtravimą pasirinktu metodu ir duomenų išvedimą į CSV failus, juos lyginti tarpusavyje nėra visiškai objektyvu. Algoritmai buvo įgyvendinti Python atvirojo kodo programavimo kalba, naudojantis skirtingų bibliotekų, skirtingomis funkcijomis, todėl laiko lyginimas tampa subjektyvus. Taip pat reiktų paminėti, jog interpoliavimo gabalais tiesine funkcija metu, signalas yra tiriamas itin mažomis atkarpomis (400 ms) ir ši tiriama atkarpa yra slenkama 80 ms žingsniu. Taip gaunama, jog tas pats taškas gali būti apskaičiuojamas 4 kartus, tai lėmė, jog morfologinių filtrų algoritmo veikimas buvo greitesnis beveik visais atvejais.

6.1. Algoritmų patikimumo vertinimas

Vertinti morfologinių filtrų ir interpoliavimo gabalais tiesine funkcija algoritmus tarpusavyje yra labai sudėtinga. Šie abu algoritmai remiasi skirtingomis teorijomis ir veikimo principais.

Morfologinių filtrų algoritmas apdoroja signalą, nufiltruodamas foninę smegenų veiklą ir išryškindamas pikus, taip suteikdamas galimybę aptikti epileptoforminius pikus, išsivedant ribą, virš kurios rastosios reikšmės būtų vadinamos pikais. Interpoliavimo gabalais tiesine funkcija algoritmas pasirinkto pločio atkarpose interpoliuoja signalą, nubrėždamas 5 tiesių rinkinį ir taip pašalindamas foninę smegenų veiklą ir palikdamas tik epileptoforminį piką.

Nepaisant skirtingų teorijų, rezultatai buvo gauti gana panašūs: abiem algoritmams buvo išvestos kelios konstantos, su kuriomis galima detektuoti epileptoforminius pikus intervalų režiuose. Morfologinių filtrų algoritme, buvo rasta detekcijos ribos koeficiento d reikšmė, lygi 9.7, tačiau 3 lentelėje pavaizduota, jog šios ribos aklas taikymas visoms encefalogramoms neduos idealių rezultatų. Patikimumas su šia konstanta svyravo tarp 35 – 96 %. Tačiau pritaikius detekcijos ribos koeficiento d reži nuo 7.76 iki 12.12, įvairiuose intervaluose skirtingų signalų rezultatai varijavo ir iki 99 %.

Interpoliavimo gabalais tiesine funkcija atveju, buvo rastas tiesių statumą apibūdinantis koeficientas m . Šis koeficientas nusako, ar tarp dviejų tiesių yra staigus kylimas į viršų ar apačią, būdingas epileptoforminiams pikams. Po testavimo su realiais pacientų duomenimis buvo gautas rezultatas : $m = 2.9$, tačiau pritaikius šį koeficientą encefalogramoms, patikimumas svyravo nuo 50% iki 99% (4 lentelė). Nužeminant rastojo koeficiento reikšmes 25 % ir pakeliant 15 %, buvo gauti režiai nuo 2.2 iki 3.3. Pritaikius šias dvi reikšmes signaluose būtų gaunami du tiesių statumą apibūdinantys parametrai, taip sudarant galimybes gydytojams ar kitiems encefalogramas tiriantiems asmenims, patiems nuspręsti, kuris dydis yra tinkamiausias.

Pateikti tikslius algoritmų jautrumo vertinimus yra sudėtinga, dėl epileptoforminių pikų savybių: skirtingiems pacientams, būdingi skirtingų charakteristikų pikai. Turimose encefalogramose buvo pastebėta, jog pacientams, kuriems aptinkama daugiau pikų, jie yra žemesni. Taip pat epileptoforminiai pikai gali būti tiek pavieniai, tiek grupėmis, todėl algoritmų įvėdinimuose reiktų įdiegti tikrinimus jų trukmei, kad keli iš eilės einantys pikai, nebūtų atpažįstami kaip vienas ligos

²DELL, procesorius: Intel(R) Core(TM) i7-6600U CPU @ 2.60GHz, 8 GB RAM, 64-bit operacinė sistema

darinys. Nužeminant abėjuose algoritmuose naudojamų koeficientų vertes tiesiog iki minimalios, padidintų pikų detekcijos jautrumą: būtų rasta žymiai didesnis skaičius pikų, tačiau detektavimo kokybės prasme, patikimumas kristų, būtų randamas didelis skaičius netikrų pikų.

Išvados ir rekomendacijos

Šis baigiamasis magistro darbas skirtas encefalogramų analizei bei epileptoforminių pikų detekcijai. Pikų detekcijos realizacija buvo įgyvendinta dviem skirtingais algoritmais. Pirmasis pagrįstas morfologinių filtrų teorija, antrasis įvykdo pikų detekciją interpoliavimu gabalais tiesine funkcija.

Morfologinių filtrų algoritme naudojamų koeficientų tikslinimui buvo sukurta vartotojo grafinė sąsaja, kuri padėjo tiksliau vizualiai nustatyti optimaliausias koeficientų reikšmes ir pateikti šias išvadas:

- struktūrinio elemento aukštis ir plotis yra vieni svarbiausių koeficientų, pikų nustatymo algoritme.
- 4.1 skyrelyje patvirtintas tinkamiausias struktūrinio elemento ilgis :

$$S_L = 4median|W|$$

Toks ilgas užtikrina, jog susidaroma struktūrinė parabolė bus platesnė, tai statistiškai vidutinis smegenų veiklos parabolės plotis ir savo pločiu primins epileptoforminį piką.

- Atkarpa morfologinių filtrų vykdymui turėtų būti tarp 4-6 sekundžių, taip susidaromas struktūrinis elementas bus optimaliausias tuo metu vyravusiems smegenų ritmams.
- 4.3 skyrelyje patvirtinta, jog parabolės struktūrinio koeficiento apskaičiavimui naudojami parametrai a_1 ir a_2 neturi būti vienodi. Pasirinktas variantas skirtumui - 3 kartai. Taip gaunamų struktūrinių parbolių aukščiai skirsis ir labiau išryškins pikus morfologinio filtravimo metu.
- Nustatyta pikų detekcijos ribos koeficiento d reikšmė, kuri turėtų svyruoti nuo 7.76 iki 12.12.

Po algoritmo, paremto interpoliavimu gabalais tiesine funkcija, įgyvendinimo galima pateikti šias išvadas:

- Algoritmas yra gana jautrus encefalogramų kokybei, taip pat epileptoforminių pikų skaičiui jose. Encefalogramose, kuriose pikų skaičius viršijo 800 pikų, tiesių koeficientų skirtumo reikšmė buvo gaunama itin maža.
- 5.1 skyrelyje buvo nustatytas optimaliausias intervalas algoritmo atlikmui: 400 ms.
- 5.2 skyrelyje gautas žingsnis, kuriuo yra slenkamas parinktas intervalas per signalą ir ieškoma optimaliausio 5 tiesių rinkinio, atitinkančio piką. Gautasis žingsnis : 80ms.
- Nustatas tiesių koeficientų modulio skirtumas $||k_1| - |k_2||$, kuris buvo gautas tarp 650 - 3000. Todėl buvo rasta šio skirtumo mediana kiekvienam signalui ir išvestas koeficientas m , kuris nusakytų kiek kartų tiesių koeficientų k_1, k_2 skirtumo modulis yra didesnis už savo medianą kiekviename signale atskirai. Taip buvo gauta m reikšmė lygi 2.2 – 3.3.

Epileptoforminių pikų analizė buvo atlikta su realių pacientų duomenimis, iš viso buvo išnagrinėtos 21 paciento encefalogramos. Abu algoritmai buvo įgyvendinti Python atvirojo kodo programavimo kalba, tačiau algoritmų palyginimas tarpusavyje yra labai sudėtingas dėl jų skirtingų veikimo specifikų ir vykdomų žingsnių. Visoms tiriamoms encefalogramoms buvo įgyvendinti

papildomi algoritmai - Furjė filtras, kurio pagalba, signaluose buvo pašalinamas 50 Hz triukšmas. Abiem algoritmų veikimo atvejais nebuvo gautas vienas pikų detektavimo vertinimas, buvo likta prie rėžių, kurių intervaluose detektuojami epileptiforminiai pikai ir paliekama vieta gydytojo ar kito encefalogramas tiriančio asmens rezultatų interpretacijai.

Ateities tyrimų planas

Galimi tolimesni darbai, pradėti magistro baigiamajame darbe:

- Epileptoforminių pikų detekcijos ribos tikslinimas morfologinių filtrų algoritmui: detekcijos ribos koeficientas d smarkiai išauga, jei pikų skaičius encefalogramoje yra mažas. Reikalingas detalesnis tyrimas, dėl kokių priežasčių, algoritmas nėra pajėgus detektuoti mažo kiekio pikų, naudojantis standartiniu nustatytu ribos detekcijos koeficientu.
- Algoritmo, paremto interpoliavimu gabalais tiesine funkcija tiriamos atkarpos bei žingsnio tikslinimas: reikalingi papildomi tyrimai, kaip nustatytas tiriamos atkarpos ilgis keičia algoritmo jautrumą pikų detekcijai.
- Algoritmo, paremto interpoliavimu gabalais tiesine funkcija, tiesių koeficientų skirtumo tikslinimas: algoritmas nedektuoja pikų, jei jų skaičius yra labai didelis encefalogramoje. Reikalingas papildomi kitų tiesių koeficientų k_3 , k_4 tikrinimų įvedimas, norint užtikrinti pilną epileptoforminių pikų detekciją.
- Morfologinių filtrų ir interpoliavimu gabalais tiesine funkcija algoritmų optimizavimas. Dabartinė algoritmų veikimo trukmė nėra optimali, laiką galima sutrumpinti lygiagrečiai encefalogramų kanalų skaičiavimą.

Literatūros šaltiniai

- [1] Selim R Benbadis. Generalized epilepsies on eeg. Medscape, 2015.
<http://emedicine.medscape.com/article/1140724-overview>.
- [2] Guangyi Chen. Automatic EEG seizure detection using dual-tree complex wavelet-fourier features. Expert Systems with Applications, 2014.
- [3] Christopher Clapham and James Nicholson. The concise oxford dictionary of mathematics. Oxford University Press Inc., New York, 2005.
- [4] Paul Garrett. Notes on first-year calculus, 2004.
http://www.math.umn.edu/~garrett/calculus/first_year/notes.pdf.
- [5] Ling Guo, Daniel Rivero, and Alejandro Pazos. Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks. Journal of Neuroscience Methods, 2010.
- [6] Jonathan J. Halford. Computerized epileptiform transient detection in the scalp electroencephalogram: Obstacles to progress and the example of computerized ECG interpretation. Clinical Neurophysiology, 2009.
- [7] Algimantas Juozapavičius, Gytis Bacevičius, and Dmitrijus Bugelskis. EEG analysis – automatic spike detection. Nonlinear Analysis: Modelling and Control, 2011.
- [8] Cheng-Wen Ko and Hsiao-Wen Chung. Automatic spike detection via an artificial neural network using raw EEG data: effects of data preparation and implications in the limitations of online recognition. Clinical Neurophysiology, 2000.
- [9] Tadas Meškauskas. Signalų analizė ir apdorojimas. Mokymo priemonė.
http://www.mif.vu.lt/~meska/SAA/Tadas_Meskauskas_-_Signalu_Analize_Ir_Apdorojimas_-_Mokymo_Priemone.pdf.
- [10] Andrius Vytautas Misiukas Misiūnas, Tadas Meškauskas, and Algimantas Juozapavičius. On the implementation and improvement of automatic EEG spike detection algorithm. Lietuvos matematikos rinkinys, 2015.
- [11] P. Modur PN and B. Rigdon. Diagnostic yield of sequential routine EEG and extended outpatient video-EEG monitoring. Clinical Neurophysiology, 2008.
- [12] Saeid Sanei and J. A. Chambers. EEG signal processing. John Willey & Sons, 2007.
- [13] William Stanley. Technical analysis and applications with matlab. Cengage Learning, 2004.
- [14] Eric W. Weisstein. Least squares fitting. A Wolfram Web Resource.
<http://mathworld.wolfram.com/LeastSquaresFitting.html>.
- [15] Webber WRS, Litt B, Lesser RP, Fisher RS, and Bankman I. Automatic EEG spike detection: what should the computer imitate? Electroencephalography and Clinical Neurophysiology, 1993.

- [16] Guanghua Xu, Jing Wang, Qing Zhang, and Junming Zhu. An automatic EEG spike detection algorithm using morphological filter. Proceeding of the 2006 IEEE International Conference on Automation Science and Engineering Shanghai, China, 2006.