Technická univerzita v Košiciach Fakulta elektrotechniky a informatiky

Algoritmy pre rozpoznávanie obrazcov pri štúdiu kozmického žiarenia ultravysokých energií v rámci JEM-EUSO experimentu

Diplomová práca

Bc. Miroslav Staroň

 $\mathbf{2013}$

Technická univerzita v Košiciach Fakulta elektrotechniky a informatiky

Algoritmy pre rozpoznávanie obrazcov pri štúdiu kozmického žiarenia ultravysokých energií v rámci JEM-EUSO experimentu

Diplomová práca

Študijný program:	Informatika
Študijný odbor:	9.2.1 Informatika
Školiace pracovisko:	Katedra počítačov a informatiky (KPI)
Školiteľ:	doc. Ing. Ján Genči, PhD.
Konzultant:	RNDr. Pavol Bobík, PhD.

Košice 2013

Bc. Miroslav Staroň

Abstrakt v SJ

JEM-EUSO experiment sa pripravuje na svoju misiu na Medzinárodnej vesmírnej stanici ISS s cieľom identifikácie zdrojov častíc kozmického žiarenia ultravysokých energií. Diplomová práca bude zameraná na štúdium signálu UV žiarenia produkovaného spŕškami sekundárnych častíc kozmického žiarenia v atmosfére Zeme a jeho odlíšenie od UV žiarenia produkovaného inými zdrojmi na nočnej strane Zeme v meraniach JEM-EUSO experimentu metódami rozpoznávania vzorov. Cieľom práce je naprogramovať modul rozpoznávania vzorov pre balík ESAF.

Kľúčové slová

Algoritmy rozpoznávania vzorov, kozmické žiarenie ultravysokých energií, JEM-EUSO experiment

Abstrakt v AJ

JEM-EUSO experiment is preparing for his mission to the International Space Station ISS. The aim is to identify the particle sources of the ultra-high energy cosmic rays. Master's thesis will focus on the study of signal UV ultra-violet radiation produced by a shower of secondary particles of cosmic rays in the Earth's atmosphere and its distinction from UV radiation, which is produced by other sources on the night side of the Earth in the measurements of JEM-EUSO experiment by methods of pattern recognition. The aim of this work is to program the module of pattern recognition for ESAF package.

Kľúčové slová v AJ

Pattern recognition algorithms, ultra high-energy cosmic ray, JEM-EUSO experiment TECHNICKÁ UNIVERZITA V KOŠICIACH

FAKULTA ELEKTROTECHNIKY A INFORMATIKY Katedra počítačov a informatiky

ZADANIE DIPLOMOVEJ PRÁCE

Študijný odbor: 9.2.1 Informatika Študijný program: Informatika

Názov práce:

Algoritmy pre rozpoznávanie obrazcov pri štúdiu kozmického žiarenia ultravysokých energií v rámci JEM-EUSO experimentu Pattern recognition algorithms for study of ultra high energy cosmic rays in

JEMEUSO experiment framework

Študent (tituly, meno, priezvisko): Bc. Miroslav Staroň

Školiteľ (tituly, meno, priezvisko):	doc. Ing. Ján Genči, PhD.
Školiace pracovisko:	Katedra počítačov a informatiky
Konzultant práce (tituly, meno, priezvisko):	RNDr. Pavol Bobík, PhD.
Pracovisko konzultanta:	Slovenská akadémia vied, Košice

Pokyny na vypracovanie diplomovej práce:

1. Oboznámiť sa so základmi fyzikálnej problematiky JEM-EUSO a experimentu so zameraním sa na rozlíšenie registrovaného UV signálu od UV pozadia nočnej oblohy 2. Oboznámiť sa so základmi fyzikálnej problematiky objavu a skúmania kozmického žiarenia ultravysokých energií (nad 4.1019 eV).

3. Preštudovať si framework a štruktúru balíka ESAF s cieľom navrhnúť a implementovať alternativny EASF modul pre pattern recognition.

4. Nájsť optimálne pattern recognition algoritmy a overiť či spoľahlivo filtrujú náhodné falošné triggrom selektované signály tvorené UV pozadím.

5. Spracovať dokumentáciu podľa pokynov vedúceho práce.

Jazyk, v ktorom sa práca vypracuje: Termín pre odovzdanie práce: Dátum zadania diplomovej práce:

slovenský 26.04.2013 31.10.2012

prof. Ing. Ján Kollár, CSc.

vedúci garantujúceho pracoviska,

prof. Ing. Liberios Vokorokos, PhD.

dekan fakulty

Čestné vyhlásenie

Vyhlasujem, že som diplomovú prácu vypracoval samostatne s použitím uvedenej odbornej literatúry.

Košice 26. 4. 2013

Vlastnoručný podpis

Poďakovanie

Touto cestou sa chcem poďakovať všetkým, ktorí mi akýmkoľvek spôsobom pomohli pri vypracovaní tejto práce. Najväčšia vďaka patrí konzultantovi diplomovej práce RNDr. Pavlovi Bobíkovi, PhD. za odborné vedenie, usmerňovanie a konzultácie, ktoré mi poskytoval pri vytváraní diplomovej práce. Rovnako sa chcem poďakovať aj svojmu vedúcemu diplomovej práce doc. Ing. Jánovi Genčimu, PhD. za odborné vedenie, usmerňovanie a poskytnutie študijných materiálov.

Obsah

	Úvo	od		1
1	JEN	M-EUS	SO experiment	3
1.1 JEM-EUSO teleskop			5	
		1.1.1	Fresnelove šošovky	6
		1.1.2	Detektory ohniskovej plochy	7
		1.1.3	Systém atmosferického monitorovania	8
		1.1.4	Kalibračný systém	9
2	Roz	zpozná	vanie vzorov	10
	2.1	Termi	nológia	11
	2.2	Návrh	n systému pre rozpoznávanie vzorov	13
		2.2.1	Zber dát	14
		2.2.2	Predspracovanie	14
		2.2.3	Extrakcia príznakov	15
		2.2.4	Výber príznakov	15
		2.2.5	Výber modelu a trénovanie	16
		2.2.6	Vyhodnotenie výkonnosti	17
3	\mathbf{Sch}	éma sj	púšťačov	18
	3.1	Prvý s	stupeň spúšťačov	19
	3.2	Druhý	ź stupeň spúšťačov	20
	3.3	Simula	ácie pozadia	22
4	Apl	ikácia	Houghovej transformácie na JEM-EUSO merania	24
	4.1	Hough	nova transformácia	24
		4.1.1	Popis metódy	25
		4.1.2	Implementácia	26
		4.1.3	Výsledky	30

	4.2	.2 Modifikovaná Houghová transformácia		
		4.2.1	Implementácia	33
		4.2.2	Výsledky	37
	4.3	Hough	ova transformácia pre pohybujúce sa vzory	39
		4.3.1	Implementácia	41
		4.3.2	Výsledky	43
5 Záver Zoznam použitej literatúry			49	
		53		
	Zoz	nam p	ríloh	57

Zoznam obrázkov

1 - 1	Schématický náčrt princípu fungovania JEM-EUSO experimentu [4]. $% \mathcal{S}_{\mathrm{e}}$.	4
1 - 2	Ilustrácia skokovitého nárastu apertúry experimentu JEM-EUSO voči	
	pozemným experimentom[4]	5
1 - 3	Schématický nákres JEM-EUSO detektora[2].	6
1 - 4	Prstencová šošovka detektora[4]	7
1 - 5	Umiestnenie fotoná sobičov na ohniskovej ploche detektora [4]	8
1 - 6	Koncepcia atmosferického monitorovania[2]	9
2 - 1	eq:Grafické znázornenie priestoru príznakov s rozhodovacími hranicami [11].	12
2 - 2	Prvky systému pre rozpoznávanie vzorov[11]	13
2 - 3	Znázornenie neurónu a MLP siete[16]	17
3 - 1	Štruktúrna schéma rôznych úrovní spúšťačov [2]	18
3 - 2	Spŕšky kozmického žiarenia znázornené na ohniskovej ploche detektora.	19
3 - 3	Bloková schéma spúšťača prvého stupňa.	20
3 - 4	Princíp spúšťača druhej úrovne[19]	21
3 - 5	Závislosť FTR od nastavenia prahových hodnôt spúšťačov	23
4 - 1	Parametre (r, θ) priamky[22]	25
4 - 2	Priestor parametrov (Houghov priestor)[20]	26
4 - 3	Rozloženie PDM v detektore[2]	27
4 - 4	Príklad generovaných matíc.	30
4 - 5	Závislosť počtu nájdených vzorov od dĺžky a priemernej hodnoty vzoru.	31
4 - 6	Nespojitý vzor.	33
4 - 7	Nespojitý vzor (pixely s hodnotu > 3)	34
4 - 8	Histogram hodnôt pixelov z LTT súborov	37
4 - 9	Histogram stredných hodnôt matíc z LTT súborov	38
4-10	Závislosť počtu nájdených vzorov od dĺžky a priemernej hodnoty vzoru.	39
4-11	lSpŕška zachytená JEM-EUSO teleskopom[21]	40
4-12	2Geometria spŕšky	41

4-13Skladanie matic.	42
4–14Závislosť počtu nájdených vzorov od dĺžky vzoru.	43
$4-15 {\rm Z}$ ávislosť počtu nájdených vzorov od dĺžky a priemernej hodnoty vzoru.	44
$4-16{\rm Z}$ ávislosť počtu nájdených vzorov s dĺžkou 2 p x v závislosti od nasta-	
venia uhlov α a β	45
$4-17 {\rm Z\acute{a}v}$ islosť počtu nájdených vzorov s dĺžkou 3 p x v závislosti od nasta-	
venia uhlov α a β	45
$4-18 {\rm Z\acute{a}v}$ islosť počtu nájdených vzorov s dĺžkou 4 p x v závislosti od nasta-	
venia uhlov α a β	46
$4-19{\rm Z}$ ávislosť počtu nájdených vzorov s dĺžkou 5 p x v závislosti od nasta-	
venia uhlov α a β	46
$4-20 {\rm Z}{\rm \acute{a}}{\rm vislosť}$ počtu nájdených vzorov s dĺžkou 6 p x v závislosti od nasta-	
venia uhlov α a β	46
$4-21 {\rm Z}{\rm \acute{a}}{\rm vislosť}$ počtu nájdených vzorov s dĺžkou 7 p x v závislosti od nasta-	
venia uhlov α a β	47
$4-22 {\rm Z}{\rm \acute{a}}{\rm vislosť}$ počtu nájdených vzorov s dĺžkou 8 p x v závislosti od nasta-	
venia uhlov α a β	47
$4-23 {\rm \check{C}asov} {\rm \check{y}}$ priebeh štyroch vybraných spŕšiek	48
$5\!-\!1$ Chyba gama-68 rekonštrukcie uhla UHECR udalostí v meraniach	
JEM-EUSO experimentu ako funkcia inklinácie primárnej častice voči	
povrchu atmosféry. Gama-68 je odchýlka 68 najbližších rekonštrukcií	
zo 100 simulovaných eventov. [1]	51
$5\!-\!2$ Závislosť počtu nájdených vzorov od ich dĺžky pre simuláciu 2500s	
meraní JEM-EUSO teleskopu	52
5-3 Aproxmácia závislosti počtu nájdených vzorov od ich dĺžky pre dl hšie	
obdobia meraní JEM-EUSO teleskopu.	52

Zoznam symbolov a skratiek

- CCB Cluster Control Board
- DNA Deoxyribonucleic acid
- EAS Extensive Air Shower
- EC Elementary Cell
- ECG Electrocardiography
- EECR Extreme Energy Cosmic Rays
- EEG Electroencephalography
- ESA European Space Agency
- EUSO Extreme Universe Space Observatory
- FPGA Field-programmable gate array
- FTR Fake Trigger Rates
- GTU Gate Time Unit 2,5 μ s
- GZK Greisen Zatsepin Kuzmin
- HTV H-II Transfer Vehicle
- ISS International Space Station
- JAXA Japan Aerospace Exploration Agency
- JEM-EUSO Japanese Experiment Module Extreme Universe Space Observatory
- LDA Linear Discriminant Analysis

- LED Light-Emitting Diode
- LIDAR Light Detection and Ranging
- LTT Linear Track Trigger
- MAPMT Multi-Anode PhotoMultiplier Tube
- MDP Mission Data Processor
- MLP Multilayer perceptron
- PCA Principal Component Analysis
- PDM Photo Detector Module
- PMT Photomultiplier tube
- PTT Persistency Track Trigger
- RBF Radial basis function
- UHECR Ultra High Energy Cosmic Ray
- UV Ultraviolet

Slovník termínov

- GZK tienenie je čakávané odfiltrovanie vysokoenergetických častíc na základe skutočnosti, že vesmír je vyplnený reliktovým žiarením (kozmické mikrovlnné pozadie CMB) – najvýznamnejším pozostatkom Veľkého tresku. Vysokoenergetické častice s najvyššou energiou sa zrážajú s CMB a strácajú energiu[19].
- **Šošovka** je homogénne a izotropné priehľadné prostredie ohraničené dvomi guľovými plochami. Ak sa lúče rovnobežné s optickou osou po prechode šošovkou zbiehajú, nazýva sa šošovka spojnou (spojkou) a ak sa rozbiehajú, nazýva sa rozptylnou (rozptylkou)[28].
- Čerenkovovo žiarenie je elektromagnetické žiarenie vznikajúce pri prechode nabitej častice prostredím rýchlosťou väčšou ako je rýchlosť svetla v tomto prostredí[3].
- Fluorescencia dusíka vzniká excitáciou atómov dusíka pozdĺž dráhy letu spŕšky sekundárnych častíc atmosférou. Pomalá deexcitácia týchto atómov vedie k sabej žiari viditeľného svetla[3].
- Mahalanobisova vzdialenosť je obecným merítkom vzdialenosti berúca v úvahu koreláciu medzi parametrami a je nazávaislá na rozsahu hodnôt parametrov. Počíta vzdialenosť medzi objektmi v systéme súradníc kde osi nemusia byť na seba kolmé. V praxi sa používa k zisteniu vzdialenosti medzi skupinami objektov[29].
- **Poissonovo rozdelenie** pravdepodobnosti sa používa na vyjadrenie rozdelenia počtu výskytov nejakého javu v určitom časovom intervale. Poissonovým rozdelením sa riadi aj počet častíc v jednotke plochy alebo objemu a pod.[30]
- Gaussovo rozdelenie pravdepodobnosti sa používa ako model v takých situáciách, keď na kolísanie náhodnej premennej pôsobí veľký počet nepatrných, navzájom nezávislých príčin[30].

Úvod

Každú sekundu narážajú na každý m^2 na Zemi tisíce nabitých častíc. Sú nazývané kozmickým žiarením. So zvyšovaním ich energie ich tok klesá a očakáva sa, že častice s energiami nad 4 . 10^{19} eV budú odfiltrované stratami spôsobenými kolíziami s mikrovlnným pozadím nachádzajúcim sa všade vo vesmíre. Prvá častica s energiou 10^{20} eV bola objavená v roku 1962 a odvtedy asi tucet ďalších v 90. rokoch. Skutočnosť, že boli spozorované častice s energiami nad 4 . 10^{19} eV nie je v súlade so súčasným stavom poznania vo fyzike a astrofyzike. Nenájdenie zdroja týchto častíc vo vzdialenosti menšej ako pár desiatok megaparsekov od slnečnej sústavy by mohlo vyvolať pochybnosti o platnosti špeciálnej teórii relativity a iných fundamentálnych fyzikálnych princípoch.

JEM-EUSO (Japanese Experiment Module - Extreme Universe Space Observatory) bol projektovaný na vyriešenie GZK(Greisen Zatsepin Kuzmin) tienenia a na identifikáciu astronomického pôvodu týchto častíc. Tento teleskop nie je smerovaný do vesmíru, ale dolu k zemskému povrchu, pretože zemská atmosféra je najrozsiahlejším doposiaľ využívaným detektorom v skúmaní pôvodu týchto častíc prichádzajúcich z vesmíru.

Diplomová práca je zameraná na štúdium signálu UV (Ultraviolet) žiarenia produkovaného spŕškami sekundárnych častíc kozmického žiarenia v atmosfére Zeme a jeho odlíšenie od UV žiarenia produkovaného inými zdrojmi na nočnej strane Zeme v meraniach JEM-EUSO experimentu metódami rozpoznávania vzorov.

Prvá kapitola sa zaoberá samotným JEM-EUSO experimentom a jeho hlavným vedeckým cieľom a to vyriešením otázky pôvodu kozmického žiarenia. V tejto kapitole je popísaný princíp fungovania JEM-EUSO teleskopu a sú uvedené jeho hlavné časti.

Pojem rozpoznávania vzorov ako vedecká disciplína je rozobratý v druhej

KPI

kapitole. Sú tu uvedené základné defínicie a popísaná terminológia problematiky rozpoznávania vzorov. Zbytok kapitoly sa venuje návrhu systému pre rozpoznávanie vzorov.

Tretia kapitola popisuje schému spúšťačov JEM-EUSO teleskopu. V kapitole je načrtnuté rozdelenie do niekoľkých úrovní a popísaný princíp fungovania každej z nich. V poslednej časti sú popísané simulácie, ktoré testujú nastevenie spúšťačov a sú zároveň zdrojom dát pre algoritmy rozpoznávania vzorov, ktoré sú analyzované v tejto práci.

V štvrtej kapitole je popísaná aplikácia Houghovej transformácie na JEM-EUSO merania. Prvú časť kapitoly tvorí popis samotného algoritmu Houghovej transformácie a jeho modifikácia. V druhej časti je rozobratá aplikácia algoritmu na JEM-EUSO merania a analýza výsledkov tejto aplikácie.

1 JEM-EUSO experiment

JEM-EUSO je UV teleskop, ktorý bude nainštalovaný na palube japonského experimentálneho modulu Medzinárodnej vesmírnej stanice ISS (International Space Station). Jeho hlavnou úlohou je monitorovať atmosféru Zeme zhora, tak aby bolo možné zaznamenať atmosferické spŕšky sekundárneho žiarenia vyvolané kozmickým žiarením ultravysokých energií (Ultra High Energy Cosmic Ray, v ďalšom texte UHECR). Základné techniky, ktoré JEM-EUSO využíva, pochádzajú z predchádzajúcej misie EUSO (Extreme Universe Space Observatory), ktorá bola z finančných a programových dôvodov zastavená v roku 2004, no uspešne absolvovala fázu A projektu[1]. Štart JEM-EUSO experimentu je naplánovaný na rok 2017 a jeho merania na ISS potrvajú minimálne 3 roky[6].

Hlavný vedecký cieľ misie je vyriešenie otázky pôvodu kozmického žiarenia ultravysokých energií. UHECR je kozmické žiarenie s energiami nad takzvaným GZK limitom (6, 10¹⁹ eV) efektívne interagujúce s mikrovlnným pozadím vesmíru. Vďaka tejto interakcii stráca energiu a na vzdialenostiach v rádoch desiatok megaparsekov postupne zníži svoju energiu pod GZK limit. Registrácia UHECR častíc na Zemi znamená, že zdroje tohto žiarenia nie sú od Solárneho systému vzdialené ďalej ako pár desiatok megaparsekov. Napriek 50 rokom výskumu tohto fenoménu však stále ide o nevyriešenú záhadu a zdroje UHECR a mechanizmy jeho urýchlovania nie sú známe[3].

JEM-EUSO experiment je spoločným projektom kozmických agentúr JAXA (Japan Aerospace Exploration Agency) a ESA (European Space Agency) a spolupracujúcich krajín. Celkový počet krajín zapojených do projektu je 13 vrátane Slovenskej republiky. Slovenskú účasť v experimente tvorí Ústav Experimentálnej Fyziky Slovenskej akadémie vied v Košiciach. Kľúčovým účastníkom JEM-EUSO projektu je Japonsko, ktoré poskytuje pre projekt základnú infraštruktúru zahŕňajúcu let nosnej rakety HII-B, kozmickú loď HTV (H-II Transfer Vehicle) a miesto pre

3

umiestnenie detektora na Medzinárodnej kozmickej stanici ISS (dok na tzv. exposure facility japonského modulu Kibo)[6].



Obrázok 1–1: Schématický náčrt princípu fungovania JEM-EUSO experimentu[4].

Častice UHECR s energiami nad GZK limit dopadajú na povrch atmosféry Zeme a produkujú v nej spŕšky sekundárneho kozmického žiarenia, ktoré excitáciou atmosférického dusíka tvoria UV žiarenie. Stopu UV žiarenia pohybujúcu sa atmosférou rýchlosťou blízkou rýchlosti svetla registruje JEM-EUSO detektor z nízkej orbity Zeme (\sim 400 km)[5].

Pri 60 stupňovej apertúre detektora pri pohľade priamo dole na Zem (v takzvanom nadir móde) tento na povrchu Zeme sleduje oblasť s rozlohou približne 140 tisíc kilometrov štvorcových 1–2. Oblasť snímkuje 400 tisíc krát za sekundu pohybujúc sa rýchlosťou 7 km/s. JEM-EUSO teleskop nie je teda smerovaný do vesmíru, ale dolu smerom k zemskému povrchu. Zatiaľ čo bežné astronomické ob-

servatórium vzhliada zo Zeme k vesmíru, JEM-EUSO pozoruje vesmír sledovaním zemskej atmosféry. Táto sa tak stane najrozsiahlejším doposiaľ využívaným detektorom v našom skúmaní pôvodu týchto nezachytiteľných častíc prichádzajúcich z vesmíru. Na obrázku 1–1 sa nachádza schématický náčrt princípu fungovania JEM-EUSO experimentu[5].



Obrázok 1–2: Ilustrácia skokovitého nárastu apertúry experimentu JEM-EUSO voči pozemným experimentom[4].

1.1 JEM-EUSO teleskop

JEM-EUSO detektor je 2,5 metrový teleskop (optická apertúra 4, $5m^2$) s vysokorýchlostnou ultrafialovou (UV v rozsahu vlnových dĺžok 300 - 400 nm) kamerou. Pozostáva z 5 tisíc fotonásobičov s celkovo 300 tisíc pixelmi, každý s rozmerom menej ako 3 x 3 mm. Jeden pixel detektora pokrýva na povrchu Zeme v nadir móde činnosti oblasť so stranou dlhou 560 metrov[7].

Detektor pozostáva zo štyroch hlavných častí: fresnelove šošovky, detektory ohniskovej plochy, systém atmosferického monitorovania a kalibračný systém[7].



Obrázok 1–3: Schématický nákres JEM-EUSO detektora[2].

1.1.1 Fresnelove šošovky

JEM-EUSO teleskop používa Fresnelove šošovky. Fresnelova šošovka je poloplochá šošovka, ktorá má kruhové drážky, ktoré eliminujú veľkú hmotnosť štandardných konvektívnych a konkávnych šošoviek. Tenkosť a ľahkosť Fresnelovej šošovky je nevyhnutnou podmienkou jej využitia vo vesmíre, pričom ponúka rovnaké optické funkcie ako hrubé a ťažké šošovky. JEM-EUSO používa dve zakrivené obojstranné Fresnelove šošovky z UV priepustného plastu a jednu mikromriežkovú Fresnelovu šošovku. Tento dizajn umožňuje najvyššiu účinnosť širokého zorného poľa. Veľkosť trojitej šošovky je 2,5 m v priemere, zložený zo stredovej 1,5 m časti a kruhového povrchu prstencových šošoviek[8].

Šošovky budú vyrobené z nového materiálu CYTOP (produkt spoločnosti Asahi Glass Company, Ltd), ktorý ešte vo vesmíne nebol použitý, no oproti bežne využívanému materiálu PMMA-000 (produkt spoločnosti Mitsubishi Rayon Company, Ltd.) má lepšie optické vlastnosti (transparentnosť, menšia zmena indexu lomu)[8].



Obrázok 1-4: Prstencová šošovka detektora[4].

1.1.2 Detektory ohniskovej plochy

Detektor ohniskovej plochy je navrhnutý tak aby zachytil stopu produkovanú spŕškou sekundárneho kozmického žiarenia v rozsahu vlnových dĺžok 300 - 400 nm. Je schopný určiť pozíciu prichádzajúcich fotonov a sledovať vývoj spŕšky v čase a priestore. Ohnisková plocha je zakrivená s priemerom 2,26 m a pokrytá multianódovými fotonásobičmi (PMTs) Hamamatsu R8900-M36. Fotonásobiče Hamamatsu majú účinnú plochu až 85%. Na celej ohniskovej ploche sa ich nachádza až 5000, pričom každý ma rozlíšenie 8 x 8 px[9]. Na obrázku 1-5 je znázornené rozloženie fotonásobičov na ohniskovej ploche.

V detektore sa nachádzajú dva stupne spúšťačov (trigrov). Prvý stupeň je implementovaný v dedikovanom FPGA čipe PDM (Photo Detector Module) modulu. Každý PDM modul je pripojený k deviatim elementárnym bunkám (EC) pozostávajúcich z 36 fotonásobičov MAPMT (Multi-Anode PhotoMultiplier Tube), pričom každý PDM modul spracúva 2304 kanálov. Výstup z každých ôsmych PDM je pripojený do jedného z 21 CCB (Cluster Control Board), ktorý prenáša informácie o pixeloch, ktoré vyhovejú podmienkam spúšťača do MDP (Mission Data Processor)[2].



Obrázok 1–5: Umiestnenie fotonásobičov na ohniskovej ploche detektora[4].

1.1.3 Systém atmosferického monitorovania

Cieľom atmosferického monitorovania je sledovanie stavu atmosféry v zornom poli JEM-EUSO teleskopu. JEM-EUSO teleskop zachytáva fluorescenciu dusíku a Čerenkovovo žiarenie emitované spŕškami sekundárneho kozmického žiarenia (Extensive Air Shower, ďalej v texte EAS) v UV pásme. Veľkosť oboch signálov zachytených JEM-EUSO detektorom zavisí od zoslabenia a rozptylu UV žiarenia v atmosfére. Správna rekonštrukcia energie EECR (Extreme Energy Cosmic Rays) a typu primárnej častice kozmického žiarenia vyžaduje informácie o absorbcii a rozptyle UV žiarenia[2].

Systém monitorovania atmosféry pozostáva z LIDARu (Light Detection and Ranging) s UV laserom a infračervenej kamery. Lidar meria optickú hrúbku a výšku mrakov, pričom táto informácia má zásadný vyznam pri rekonštrukcii udalosti, keď sa z parametrov spŕšky určuje energia a smer príchodu primárnej UHECR častice. Infračervená kamera meria teplotu v zornom poli teleskopu s presnosťou lepšou ako 3 K[1].

FEI



Obrázok 1-6: Koncepcia atmosferického monitorovania[2].

Princíp atmosferického monitorovania je znázorený na obrázku 1–6. Rozloženie mrakov v zornom poli bude nepretržite monitorovat infračervená kamera, ktorá bude tiež merať výšku hornej vrstvy opticky hustých mrakov. LIDAR určí podrobné vlastnosti atmosféry (rozptyl a strata UV žiarenia) v mieste, kde bola zachytená spŕška. Na odvodenie parametrov dôležitých pre modelovanie vlastností prenosu a zoslabenia UV žiarenia atmosférou, potrebných pre analýzu dát z LI-DARu, kalibráciu infračervenej kamery a pre modelovanie vývoja EAS v atmosfére budú použíte real-time modely atmosféry[2].

1.1.4 Kalibračný systém

Kalibračný systém pozostáva z troch LED (Light-Emitting Diode) diód s rôznou vlnovou dĺžkou (300 - 500 nm), ktoré produkujú difúzne UV žiarenie pre kalibráciu zisku a účinnosti detekcie detektora ohniskovej plochy. Okrem toho búdu na vybraných pozíciách na zemskom povrchu rozmiestnené xenónové smerové svetlá a lidarové stanice produkujúce známe množstvo UV fotónov. Detegovaním časti týchto fotónov JEM-EUSO detektorom sa určí absorbcia UV žiarenia v atmosfére a realizuje sa priebežná kalibrácia teleskopu počas jeho meraní na ISS[6].

2 Rozpoznávanie vzorov

Rozpoznávanie vzorov je vedecká disciplína, ktorej cieľom je klasifikácia objektov do niekoľkých ktegórií alebo tried. V závislosti od použitia, tieto objekty môžu byť obrázky, priebeh signálu alebo akékoľvek meranie, ktoré potrebuje byť klasifikované[10].

Základným problémom v automatizovanom rozpoznávaní a rozhodovaní je skutočnosť, že problém rozpoznania, ktorý je jednoduchý pre 5 ročné dieťa, môže byť ťažko riešiteľný pre počítače[11]. Ako príklad môžeme uviesť rozpoznanie pohlavia osôb na fotografii. Je mnoho príznakov, ako napríklad dĺžka vlasov, pomer výšky a váhy, telesné krivky, výraz v tvári a pod., ktoré by mohli použiť počítače na rozoznanie pohlavia osôb z fotografií. No nie je tažké si uvedomiť, že pri kombinácii týchto príznakov nemusí byť určenie pohlavia správne. Hoci môžeme ľahko identifikovať každú osobu ako muža alebo ženu, nie je ľahké si uvedomiť ako sme k tomuto záveru prišli, alebo presnejšie, ktoré príznaky sme k tomu použili.

Rozpoznávanie vzorov pokrýva široké spektrum aplikácií, ako napríklad rozpoznávanie reči, identifikácia hovoriacich, rozpoznávanie rukou písaného textu, diaľkový topografický prieskum, detekcia poruchy systému na základe dát zo senzorov alebo napríklad rozhodovanie o pôžičkách na základe finančných hlásení a mnoho ďalších. V poslednej dobe sa rozpoznávanie vzorov čoraz častejšie uplatňuje aj v biomedicínskych aplikáciách ako napríklad identifikácia DNA (Deoxyribonucleic acid) sekvencií, automatizované ECG (Electrocardiography) a EEG (Electroencephalography), automatizovaná analyzá snímkov a pod. Všetky tieto aplikácie majú spoločný menovateľ a to automatizovanú klasifikáciu alebo rozhodovanie na základe pozorovaných parametrov ako signál, obraz alebo všeobecne vzorov získaných kombináciou niekoľkých pozorovaní alebo meraní[11].

2.1 Terminológia

Premenné, o ktorých sa predpokladá, že su nositeľmi charakteristických informácií o objekte, ktorý ma byť identifikovaný sa nazývajú príznaky. Tieto premenné pochádzajú zvyčajne z meraní alebo pozorovaní objektu. Súbor príznakov s veľkosťou d usporiadaný do d-rozmerného stĺpcového vektora sa nazýva príznakový vektor, zvyčajne sa označuje ako x a predstavuje popis objektu, ktorý chceme identifikovať. D-rozmerný priestor, v ktorom sa tento vektor nachádza nazývame priestor príznakov. D-rozmerný vektor v d-rozmernom priestore predstavuje bod v tomto priestore príznakov. Kategória, do ktorej patrí daný objekt sa nazýva trieda a zvyčajne sa označuje ako ω . Súbor príznakov použitých pri rozhodovaní o správnej triede objektu sa nazýva vzor[10].

Cieľom systému pre rozpoznávanie vzorov je teda stanoviť správne označenie zodpovedajúce danému vektoru príznakov na základe predchádzajúcich znalostí získanych trénovaním. Trénovanie je proces, ktorým sa systém pre rozpoznávanie vzorov učí správne mapovanie medzi vektorom príznakov a zodpovedajúcim označením. Toto mapovanie vytvára rozhodovacie hranice v d-rozmernom priestore príznakov, ktoré oddeľujú vzory rôznych tried. Môžeme teda konštatovať, že cielom algoritmu rozpoznávania vzorov je stanovenie rozhodovacích hraníc, ktoré vo všeobecnosti predstavujú nelineárne funkcie. V dôsledku čoho, môže byť problém rozpoznávania vzorov pretransformovaný ako problém aproximácie funkcí. Rozpoznávanie vzorov je všeobecne rozdelené do kategórií podľa procedúry učenia, použitej na generovanie výstupnej hodnoty[10].

Kontrolované učenie predpokladá, že bola stanovená množina trénovacích dát, skladajúca sa zo sady príkladov, ktoré boli riadne označené a bola pre ne aj stanovená správna výstupná hodnota. Procedúra učenia generuje model, ktorý sa pokúša splniť dva, niekedy protichodné ciele: v čo najväčšej možnej miere reprezentovať trénovaciu množinu a zaroveň byť tak všeobecným aby pokryl aj nové údaje.



Obrázok 2–1: Grafické znázornenie priestoru príznakov s rozhodovacími hranicami[11].

Nekontrolované učenie na druhej strane predpokladá, trénovaciu množinu, ktorá nebola ručne označená a snaží sa hľadať súvisiace vzory v dátach, ktoré potom môžu byť použité na správne určenie výstupnej hodnoty pre nové dáta.

Kombinácia predchádzajúcich dvoch typov predsavuje čiastočne kontrolované učenie, ktoré používa kombináciu označených a neoznačených údajov (zvyčajne malý súbor dát označených v kombinácii s veľkým množstvom neoznačených dát).

Algoritmy pre rozpoznávanie vzorov môžeme rozdeliť podľa typu výstupného označenia, podľa použitej procedúry učenia a podľa toho či je algoritmus štatistickej povahy. Štatistické algoritmy môžu byť generatívne alebo diskriminačné.

Kvantitatívne vyjadrenie ceny chyby pri identifikácii sa nazýva nákladová funkcia (cost function). Algoritmus rozpoznávania vzorov je špeciálne navrhnutý aby minimalizoval túto funkciu. Obvykle je hodnotiaca funkcia vyjadrená ako priemerná kvadratická chyba medzi skutočnými a predpokladanými číselne vyjadrenými označeniami. Systém pre rozpoznávanie vzorov, ktorý ma pomocou procedúry učenia nastavené parametre tak, že hranice rozhodovania sú správne určené a hodnotiaca funkcia je minimalizovaná nazývame klasifikátor alebo formálnejšie ako model[11].

Pri hodnotení výkonnosti systémov sú často používané dva parametre. Schopnosť alebo výkonnosť klasifikátora správne určiť triedu trénovacích dát sa na-

12

zýva výkonnosť pri trénovaní (učení). Výkonnosť pri trénovaní sa používa na stanovenie ako dlho bude učenie prebiehať, alebo ako kvalitne sa systém naučil rozoznávať trénovacie dáta. Výkonnosť pri učení vo všeobecnosti nie je dobrým ukazovateľom všeobecnej výkonnosti, ktorá reprezentuje schopnosť alebo výkonnosť klasifikátora určiť triedu vzorov, ktoré ešte nikdy nevidel[11].

2.2 Návrh systému pre rozpoznávanie vzorov

Návrh systému pre rozpoznávanie vzorov pozostáva z niekoľkých krokov, ktoré sú zobrazené na obrázku 2-2.



Obrázok 2-2: Prvky systému pre rozpoznávanie vzorov[11].

2.2.1 Zber dát

Okrem výberu vhodného klasifikačného algoritmu, je jednou z najdôležitejších požiadaviek na návrh úspešného systému adekvátne množstvo reprezentatívnych dát, používaných na trénovane a testovanie systému. Neexistuje žiadne pravidlo, ktoré hovorí, aké množstvo dát je potrebné, avšak všeobecne platí, že tréningová množina by mala obsahovať 10 krát viacej inštancií ako je počet nastaviteľných parametrov klasifikátora[12] alebo 10 krát viacej inštancií ako ako tried a príznakov[13].

Reprezentatívne dáta na druhej strane zabezpečia, že všetky významné variácie, ktoré ma systém vedieť rozpoznať budú zahrnuté v trénovacej a testovacej množine. Podmienka reprezentatívnosti dát sa často ťažko dosahuje. Najvhodnejšie je použiť taký rozsah aký použil dizajnér pri výbere senzorov, ktoré budú poskytovať dáta pre spracovanie[11].

2.2.2 Predspracovanie

Predspracovanie je dôležitým no často prehliadaným krokom v procese navrhovania. Cieľom predspracovania je dosiahnúť čo najlepšiu kvalitu dát odstránením šumu. Na odstránenie šumu je možné použiť rôzne filtračné techniky, pri použití ktorých je ale potrebné poznať spektrum šumu.

Predspracovanie môže zahŕňať aj odstránenie odľahlých bodov. Pre menšie problémy, kde $d \leq 3$, môžeme dáta vizualizovať a tak zistiť prítomnosť odľahlých bodov. Pre rozsiahlejšie problémy, kde d > 3, môže byť použitá Mahalanobisova vzdialenosť a to najmä v prípadoch dát s Gaussovou distribúciou alebo s distribúciou podobnou Gaussovej. V tomto prípade sa vypočíta Mahalanobisova vzdialenosť každého dátového bodu z tréningovej množiny pre danú triedu. Ak je vzdialenosť väčšia ako stanovená hranica (najčastejšie štandardná dochýlka), môžeme tento bod označiť ako odľahlý[13].

FEI

2.2.3 Extrakcia príznakov

Extrakcia priznakov rovnako ako výber príznakov sú v konečnom dôsledku prostupy, ktoré umožňujú znížiť počet dimenzií. Stručne povedané, cieľom extrakcie príznakov je nájsť pokiaľ možno čo najmenší počet príznakov, ktoré sú obzvlášť charakteristické a informatívne a zároveň invariantné voči irelevantným transformáciám[11].

Malá ale informatívna množina príznakov výrazne znižuje zložitosť klasifikačného algoritmu a taktiež aj jeho časovú a pamäťovú náročnosť. Nepriaznivý efekt veľkého počtu príznakov je v komunite, ktorá sa zaoberá rozpoznávaním vzorov dobre známy a nazýva sa *prekliatie dimenzionality*. Preto je dôležité aby bol počet príznakov minimalizovaný, tak ako je to len možné pri súčasnom zachovaní schopnosti identifikovať vzor[14].

Na extrakciu príznakov sa najčastejšie používajú matematické transformácie nad získanými dátami. Niektoré z najviac používaných transformácií sú analýza hlavných komponentov (PCA - Principal Component Analysis) a lineárna dikriminačná anlýza (LDA - Linear Discriminant Analysis)[14].

Hoci sú PCA a LDA veľmi často používané, nie sú nevyhnutne tie najlepšie pre všetky aplikácie. V závislosti od aplikačnej oblasti môžeme lepšie výsledky dosiahnúť pomocou transformácie Fourierovho typu ako napríklad disktrétna alebo rýchla Fourierova transformácia [11].

2.2.4 Výber príznakov

Pod výberom príznakov sa rozumie výber podmnožiny veľkosti m z už existujúcej množiny príznakov, ktorá bola identifikovaná v predchádzajúcom kroku. Kľúčovou otázkou je, ktorá podmnožina poskytuje najlepsie možnosti rozlíšenia dát. Najčastejšie kritérium pre posúdenie rozlišovacej schopnosti vybranej podmnožiny príznakov je všeobecná výkonnosť klasifikátora, ktorý bol trénovaný na tejto podmnožine. Pri výbere správnej podmnožiny príznakov je potrebné zohľadniť aj typ použitého klasifikátora, pretože podmnožina príznakov vhodná pre neurónovú sieť nemusí byť vhodná pre klasifikátor založený na rozhodovacom strome[14].

Konceptuálne triviálnym riešením by bolo osobitne vytrénovať klasifikátor pre každú podmnožinu príznakov a vybrať tú, pri ktorej dosahuje klasifikátor najlepšiu všeobecnú výkonnosť. Takéto hľadanie tej najlepšej podmnožiny je však príliš výpočtovo náročné aj pre relatívne malý počet príznakov. Je to zapríčinené tým, že počet podskupín, ktoré majú byť vyhodnotené je kombinačným číslom z počtu príznakov a veľkosti hľadanej podmnožiny[11].

$$C(d,m) = \frac{d!}{m!(d-m)!}$$
(2.1)

Predchádzajúce vyjadrenie je platné, ak nás zaujímajú len podmnožiny s počtom prvkov presne m. Ak chceme zistiť celkový počet podmnožín v prípade, že nás zaujímajú všetky podmnožiny menšie nanajvýš rovné m, musíme použiť vzorec 2.2.

$$\sum_{i=1}^{m} C(d,i) \tag{2.2}$$

Z uvedeného vyplýva, že na nájdenie najvhodnejšej podmnožiny bude výhodnejšie použiť niektorý z už existujúcich algoritmov pre vyhľadávanie, ako napríklad prehľadávanie do hĺbky, prehľadávanie do šírky, metóda vetiev a hraníc prípadne gradientný algoritmus.

2.2.5 Výber modelu a trénovanie

Ako už bolo spomenuté, o klasifikácii môžeme uvažovať ako o probléme aproximácie funkcií a teda nájdení takej funkcie, ktorá by mapovala d-dimenzionálny vstup na vhodne kódovaný údaj o triede. Akonáhle začneme o klasifikácii uvažovať ako o probléme aproximácie funkcií môžeme použiť množstvo matematických nástrojov ako napríklad optimalizačné algoritmy. Klasifikátory a teda algoritmy rozpoznávania vzorov možeme rozdeliť na základe použitého prístupu do dvoch skupín:

- štatistické,
- neurónové siete.

Tieto dve skupiny algoritmov medzi sebou úzko súvisia a je možné dokázať, že pre algoritmus z triedy štatistických existuje ekvivalentná néurónová sieť a opačne[15]. Medzi najpoužívanejšie neurónové sietev rámci rozpoznávania vzorov patria MLP (Multilayer perceptron) a RBF (Radial basis function) siete[11]. Na obrázku 2-3 je zobrazený neurón a MLP sieť.



Obrázok 2–3: Znázornenie neurónu a MLP siete[16].

2.2.6 Vyhodnotenie výkonnosti

Akonáhle bol zvolený model (klasifikátor) a bol natrénovaný na trénovacej množine je potrebné vyhodnotiť jeho skutočný výkon pomocou dát, ktoré neboli použité pri trénovaní a klasifikátor ich teda ešte nevidel.

Jednou z najpoužívanejších metód pre hodnotenie všeobecnej výkonnosti je rozdeliť testovacie dáta do dvoch skupín, pričom prvá skupina sa používa na vlastné trénovanie a druhá na testovanie výkonu. Výkonnosť algoritmu na tejto druhej množine môžeme potom použiť na odhad jeho skutočnej výkonnosti na reálnych dátach[11].

3 Schéma spúšťačov

Hlavnou úlohou spúšťačov je spomedzi šumu tvoreného pozadím detegovať výskyt signálu, ktorý môže potencionálne znamenať prítomnosť spŕšky kozmického žiarenia v mieste aktuálneho pozorovania. Vzhľadom na veľký počet pixelov ohniskovej plochy bola pre JEM-EUSO navrhnutá viacúrovňová schéma spúšťačov[17].



Obrázok 3–1: Štruktúrna schéma rôznych úrovní spúšťačov[2].

FEI

Navrhnutá dvojúrovňová schéma spúšťačov tak ako je zobrazená na obrázku 3–1 vychádza z rozdelenia ohniskovej plochy na menšie časti, takzvané PDM, ktoré sú dostatočne veľké, aby mohli zachytiť podstatnú časť stopy spŕšky. Systém spúšťania funguje na základe štatistických vlastností prichádzajúceho toku fotónov s cieľom nájsť zaujímavé prípady rušené pozadím, pričom berie do úvahy koreláciu ich polohy a času[17].



Obrázok 3–2: Spŕšky kozmického žiarenia znázornené na ohniskovej ploche detektora.

3.1 Prvý stupeň spúšťačov

V rámci prvého stupňa spúšťačov LTT (Linear Track Trigger) je implementovaný nultý stupeň, ktorého úlohou je odfiltrovať pozadie, ktoré nemá nič spoločné s prichádzajúcimi UV fotónmi. Nultý stupeň odfiltrováva elektronický šum produkovaný v posledných stupňoch fotoelektrónového zosilnenia signálu v emisnej eletrónke a šumu, ktorý vnáša elektronika. Pre každý pixel elementárnej bunky (EC) sú digitalizované anódové pulzy v rámci jedného GTU (Gate Time Unit 2,5 μ s) spočítavané v čítači *C1* a porovnané s určenou prahovou hodnotou *N*. V každom GTU je čítač *C1* znulovaný. Ak počet impulzov v čítači *C1* prekročí preddefinovanú prahovú hodnotu sú všetky nasledujúce impulzy prenesené do druhého čítača C2 a je aktivovaný signál L, ktorý značí, že daný pixel bol aktívny. Všetky signály L sú privedené do logického súčtového člena (OR)[2].

Výstup O z logického člena OR, do ktorého boli privedené signály L je privedený na vstup čítača C3, ktorý je spoločný pre celú elementárnu bunku (Elementary Cell, skratka EC). Ten je inkrementovaný pri každom GTU vždy ak je aktívny výstup O inak je zresetovaný. Ak počet impulzov v čítači C3 prekročí prahovú hodnotu P, je signál privedený do sčítačky A, ktorá obsahuje súčet čítačov C2 zoskupených pixelov 2x2 (alebo 3x3). Výsledný súčet je porovnaný s preddefinovanou prahovou hodnotou S, ktorá predstavuje vyžadovaný počet anódových impulzov v zoskupení pixelov 2x2 (alebo 3x3). Ak je súčet väčší je vygenerovaný EC spúšťač. Na obrázku 3-1 je znázornené blokové zapojenie spúšťača prvého stupňa[2].



Obrázok 3-3: Bloková schéma spúšťača prvého stupňa.

3.2 Druhý stupeň spúšťačov

Spúšťač druhej úrovne PTT (Persistency Track Trigger) oddeľuje unikátny vzor spŕšky od pozadia sledovaním pohybu bodu vytváreného spŕškou v preddefinovanom časovom okne v rámci jedného modulu PDM. Sledovanie začína v mieste kde bol aktivovaný spúšťač prvej úrovne a to vytvorením oblasti s rozmermi 3 x 3 px, ktorá je posúvaná pri každom GTU pozdĺž preddefinovaných smerov. Celkový počet fotónov pozdĺž každého smeru je integrovaný pri každom GTU sumáciou počtu pixelov v danej oblasti za predpokladu, že ich počet je v danom GTU väčší ako prahová hodnota S. Integrácia sa vykoná pre 15 po sebe nasledujúcich GTU. Spúštač je aktivovaný ak výsledok integrácie pozdĺž daného smeru prekročí stanovenú prahovú hodnotu[18].



Obrázok 3–4: Princíp spúšťača druhej úrovne[19].

3.3 Simulácie pozadia

Cieľom simulácií pozadia je vyhodnotiť počty aktivovaných spúšťačov, takzvané FTR (Fake Trigger Rates) na generovanom pozadí. UV pozadie zachytené JEM-EUSO má náhodnú distribúciu Poissonovho charakteru.

Testovanie spúšťačov prebieha na generovanom pozadí s Poissonovskou distribúciou priemerne 500 fotónov $(m^{-2}s^{-1}sr^{-1})$ čo je približne 2,1 fotónu na jeden pixel v jednom GTU. Pri testovaní je simulované jedno PDM, pričom sú menené prahové hodnoty spúšťačov. Pri simulácii sa generuje obrovské množstvo dát, ktoré je následne použité pri testovaní rozpoznávania vzorov.

Pre jedno PDM bolo nasimulovaných 10 miliárd GTU, čo znamená 2500 sekundové meranie pozadia na jednom PDM module. Po aplikovaní schémy spúšťačov bol pôvodný počet GTU zredukovaný takmer na jednu stotisícinu. Simulácie prezentované v tejto práci boli vykonané počas viacerých mesiacov s využitím približne 200 CPU jadier JEM-EUSO výpočtového clustra.

Nasimulovanie kompletného merania celého detektora, to znamená 137 PDM modulov počas predpokladaného 3 ročného obdobia zberu dát je mimo dosiahnuteľných výpočtových kapacít JEM-EUSO kolaborácie. Simulácia kompletného merania presahuje použitých 2500 sekúnd pre jeden PDM modul o 6 dekadických rádov v dobe trvania výpočtu aj úložnej kapacite.

Na obrázku 3–5 je znázornená závislosť počtu výskytov falošných spúšťačov od nastavenia ich prahových hodnôt pre konfiguráciu 36 x 36 px (M36), ktorá je použitá aj pri testovaní rozpoznávania vzorov v tejto práci.



Obrázok 3–5: Závislosť FTR od nastavenia prahových hodnôt spúšťačov.
4 Aplikácia Houghovej transformácie na JEM-EUSO merania

JEM-EUSO teleskop deteguje spŕšku ako stopu na ohniskovej ploche, korej geometria, intenzita a časovanie umožňujú rekonštrukciu smeru a maxima spŕšky a teda aj energiu a smer primárnej častice. Hlavnou úlohou rozpoznávania vzorov je teda extrakcia stopy spŕšky z pozadia a určenie jej časových a priestorových parametrov, ktoré sú následne použité pri rekonštrukcii jej smeru a energie primárnej častice[21].

V rámci JEM-EUSO experimentu sa uvažuje o použití dvoch algoritmoch pre rozpoznávanie vzorov:

- klastering,
- Houghova transformácia

V tejto práci sa budeme zaoberať použitím Houghovej transformácie.

4.1 Houghova transformácia

Houghová transformácia je technika extrakcie charakteristických znakov využívaná pri analýze obrazu, počítačovom videní a digitálnom spracovaní obrazu. Cieľom metódy je nájsť nedokonalé inštancie objektov určitej triedy pomocou hlasovacej procedúry. Toto hlasovanie sa vykonáva v priestore parametrov, z ktorého sú kandidáti objektov získaní ako lokálne maximá v tzv. akumulačnom priestore, ktorý je zostrojený počas výpočtu[21].

4.1.1 Popis metódy

V obrazovom priestore môže byť priamka popísaná ako y = kx+q a graficky znázornená pre každú dvojicu bodov (x, y). Hlavnou myšlienkou Houghovej transformácie je pohľad na priamku z hľadiska jej parametrov k a q. Na tomto základe, môže byť priamka y = kx+q reprezentovaná ako jeden bod (k, q) v priestore parametrov. Avšak tejto reprezentácii bráni fakt, že hodnoty parametrov k a q môžu byť neobmedzené. Z výpočtových dôvodov je lepšie použiť inú dvojicu parametrov, označených ako r a θ . Tieto parametre sú polárne súradnice. Parameter r predstavuje najmenšiu vzdialenosť medzi priamkou a východiskovým bodom a parameter θ predstavuje uhol spojnice medzi východiskovým bodom a najbližším bodom priamky[22].



Obrázok 4–1: Parametre (r, θ) priamky[22].

Pomocou tejto parametrizácie môžeme rovnicu priamky prepísať ako:

$$y = \left(-\frac{\cos\theta}{\sin\theta}\right)x + \left(\frac{r}{\sin\theta}\right) \tag{4.1}$$

Po úprave dostávame:

$$r = x \, \cos\theta \, + \, y \, \sin\theta \tag{4.2}$$

Každú priamku v danom obrázku je teda možné asociovať s dvojicou (r, θ). Priestor (r, θ) sa niekdy nazýva Houghov priestor pre množinu priamiek v dvojrozmernom priestore[20].

Pre ľubovoľný bod z obrázku môžeme v priestore parametrov vykresliť sínusoidnú krivku, ktorej body (r, θ) zodpovedajú priamkám prechádzajúcim daným bodom. V prípade, že sa dve krivky v niektorom z bodov prekryjú, zodpovedá tento priamke v pôvodnom obrázku, prechádzajúcej cez tieto dva body. Vo všeobecnosti môžeme povedať, že množina bodov v obrázku zodpovedajúca priamke vyprodukuje v priestore parametrov sínusoidy, ktoré sa pretínajú v bode zodpovedajúcom tejto priamke[20].



Obrázok 4–2: Priestor parametrov (Houghov priestor)[20].

4.1.2 Implementácia

Z popisu algoritmu vyplýva, že problém detekcie bodov ležiacich na jednej priamke môžeme previesť na problém hľadania súbežných kriviek v houghovom priestore. Na detekciu sa používa pole, nazývané aj akumulátor. Dané pole ma toľko rozmerov koľko je neznámych parametrov. V našom prípade sú to dva parametre, r a θ .

V našej implementácii vyhľadávame vzory v maticiach 8 x 8, kde každý prvok (pixel) môže nadobudnúť hodnoty 0 až 7. Hodnoty prvkov matice sú ge-

nerované náhodne s uniformným rozdelením náhodných čísel v zadanom rozsahu generovaných hodnôt. Rozmer matice 8x8 je totožný s rozmerom fotomultiplikátora Hamamatsu M64 používaného pri meraniach v JEM-EUSO experimente označovaného v štruktúre detektora ako PMT. Fotomultiplikátory PMT sa zoskupujú do väčších celkov zvaných PDM, ktoré obsahujú 36 PMT. Celý detektor pozostáva zo 137 PDM modulov.



Obrázok 4–3: Rozloženie PDM v detektore[2].

V prvom kroku prejdeme všetky body spracovávaného snímku a zostrojíme houghov priestor pre všetky pixely jednej matice 8 x 8, tak ako sme to popísali v časti Popis metódy. Do výpočtu zahŕňame len tie body, ktorých hodnota je vyššia ako zvolená prahová hodnota, ktorá je pre účely nášho experimentu stanovená na hodnotu 3.

1. Výpis: Zostrojenie Houghovho priestoru.

```
1 FOR x=1,8 DO BEGIN
```

```
<sup>2</sup> FOR y=1,8 DO BEGIN
```

```
IF (matica [x-1,y-1] \text{ GT } 3) THEN BEGIN
3
         pixel = matica [x-1, y-1]
4
        FOR t=0,11 DO BEGIN
5
       theta = t * 15 * 3.14159 / 180
6
       ro = x * cos(theta) + y * sin(theta)
7
       hough_space[x-1,y-1,t] = ro
        ENDFOR
g
      ENDIF
    ENDFOR
11
12 ENDFOR
```

Prahová hodnota súvisí s fyzikálnou situáciou monitorovanou JEM-EUSO detektorom. Detektor sníma intenzitu UV žiarenia na nočnej strane Zeme, najmä počas bezmesačných nocí. Neustále tak vidí UV pozadie tvorené takzvaným airglow žiarením, zodiakálnym svetlom a odrazeným svetlom hviezd. Tento UV šum produkuje intenzity ktoré sú nižšie ako intenzity UV stôp spôsobených spŕškou sekundárneho žiarenia vyvolaného UHECR udalosťou. Pozadie je potrebné filtrovať z dôvodu limitovaného dátového toku z ISS prideleného JEM-EUSO experimentu. Zaznamenávajú sa tak len pixely s hodnotou presahujúcou stanovenú prahovú hodnotu.

V následujúcom kroku prechádzame zostrojený houghov priestor a pre každý smer (uhol) zostrojíme akumulátor. V tomto sa naša implementácia odlišuje od klasickej, keďže používame jednorozmerný akumulátor no pre každý smer ho inicializujeme nanovo. Následne sa podľa parametru r inkrementuje hodnota akumulátora.

```
2. Výpis: Naplnenie akumulátora.
```

```
1 accumulator = fltarr(400)
2 accumulator_pixels = fltarr(400)
3 FOR x=0,7 DO BEGIN
4 FOR y=0,7 DO BEGIN
5 ro = hough_space[x,y,t]
6 IF (ro GT 0) THEN BEGIN
```

```
7  kq = round(ro * 10) + 200
8  accumulator[kq]++;
9  accumulator_pixels[kq] += matica[x,y]
10  ENDIF
11  ENDFOR
12  ENDFOR
```

V poslednom kroku hľadáme maximum v akumulátore, ktoré ak spĺňa ďalšie zvolené prahové hodnoty zodpovedá priamke v spracovávanom snímku.

```
1 \max = 0
  sum_max = 0
2
  FOR k=1,399 DO BEGIN
3
     IF (accumulator [k] GT max) THEN BEGIN
       \max = \operatorname{accumulator}[k]
       sum_max = accumulator_pixels[k]
6
    ENDIF
  ENDFOR
8
  IF ((max GT 6) && ((sum_max/max) GT 6.9)) THEN BEGIN
9
    PRINT, "Nasiel som vzor"
11 ENDIF
```

V našej implementácii hľadáme také priamky, ktoré pozostávajú aspoň zo 7 bodov, ktorých priemerná hodnota je viac ako 6.9. Tu prichádzame k dvom parametrom selekcie vzorov v houghovom priestore. Prvým je dĺžka vzoru, druhým priemerná hodnota pixelov vybraného vzoru. Dĺžka vzoru je parameter určujúci počet pixelov s hodnotou nad určeným prahom (tu 3) v smere vzoru. Priemerná hodnota vzoru je suma hodnôt všetkých pixelov vzoru predelená dĺžkou vzoru.

4.1.3 Výsledky

Vzhľadom na využitie metódy pre JEM-EUSO experiment nás nezaujíma len vyhľadávanie vzorov v obrázku ale aj možnosť vzniku falošných vzorov v skúmaných maticiach. Preto netestujeme len či algoritmus nájde vzor v maticiach, ale aj to aká je pravdepodobnosť, že nájde falošný vzor, ktorý vznikol z obyčajného šumu.



Obrázok 4–4: Príklad generovaných matíc.

Dôvodom je sústredenie sa na určenie možnosti vzniku falošných vzorov v náhodne generovanom signále. UV pozadie registrované JEM-EUSO detektorom je náhodného charakteru. Je potrebné zistiť či náhodne procesy, ktoré ho vytvárajú neprodukujú občas takzvane falošné vzory, ktoré by sa mohli mylne vyhodnotiť ako UHECR udalosti.

Aplikácia Houghovej metódy bola pred aplikáciou na matice s generovaným šumom kontrolovaná vložením malého množstva vzorov do množiny matíc s nagenerovaným šumom. Metóda spoľahlivo detegovala umelé vzory.

Na obrázku 4-5 je vizualizácia závislosti strednej hodnoty pixelov vzoru od počtu zaregistrovaných vzorov, pre rôzne dĺžky vzoru (4 až 8). Počet detekovaných vzorov na osi y určuje počet vzorov s dĺžkou 4 (červená) až 8 (tmavomodrá) s priemernými hodnotami určenými hodnotami na osi x v 10 miliónoch nagenerovaných maticiach.



Obrázok 4-5: Závislosť počtu nájdených vzorov od dĺžky a priemernej hodnoty vzoru.

Z grafu na obrázku 4–5 môžeme napríklad odčítať, že v prípade ak nagenerujeme 10 miliónov matíc s rozmermi 8 x 8 px, tak približne 20 z nich bude obsahovať falošný vzor dlhý 8 pixelov s priemernou hodnotou 7 (tj. všetky pixely vzoru majú maximálnu hodnotu). Túto hodnotu môžeme overiť aj jednoduchou úvahou. Pravdepdobnosť, že pixel matice nadobudne jednu konkrétnu hodnotu je 1/8, keďže nadobúda 8 celočíselných hodnôt. Akákoľvek formácia 8 pixelov tj. aj priamy vzor dlhý 8 pixelov s rovnakou hodnotou pixelov vznikne s pravdepodobnostou $(1/8)^8 = 5,96 \cdot 10^{-8}$. Takýchto vzorov môžeme na matici 8x8 rozpoznať 32. Pri pokusoch sme ale zistili, že klasická Houghová transformácia nevie rozlíšiť súvislé a nesúvislé (prerušné) vzory. To znamená, že ak klasická houghova transformácia zistila, že v matici sa nachádza vzor s dĺžkou napríklad 5 pixelov, neznamenalo to, že tieto pixely tvorili vo vstupnej matici jednu úsečku, ale len to, že tieto pixely sa nachádzajú na jednej priamke, ktorá je daná parametrami (r, θ). Tiež sme zistili, že ak sa na jednej priamke nachádza niekoľko vzorov rôžnych dĺžok, pomocou Houghovej transformácie tieto nievieme rozlíšiť.

4.2 Modifikovaná Houghová transformácia

Pre potreby použitia Houghovej transformácie v rámci JEM-EUSO experimentu je potrebné algoritmus upraviť tak, aby vedel rozlíšiť spojité vzory od nespojitých. Na obrázku 4-6 môžeme vidieť vizualizovanú maticu hodnôt s rozmermy 36 x 36, do ktorej bol zámerne vložený jasný nepsojitý vzor. Do riadku č. 15 sme vložili dva vzory, jeden s dĺžkou 8 a druhý s dĺžkou 6.

Na obrázkku 4–7 je zobrazená tá istá matica po odfiltrovaní pixelov, ktorých hodnota bola pod stanovenou prahovou hodnotou. Môžeme vidieť, že na riadku č. 15 sa nachádza 18 pixelov nad stanovenou prahovou hodnotou, ktoré po aplikovaní Houghovej transformácie tvoria jeden vzor s dĺžkou 18 a priemernou hodnotou 6,4. Pre JEM-EUSO je potrebné upraviť algoritmus tak aby dokázal nájsť obidva vložené vzory. Jeden dĺžky 8 s priemernou hodnotou 7 a druhý s dĺžkou tiež 8 a priemernou hodnotou 6,4. Druhý vzor vznikol z vloženého vzoru a susedných pixelov s hodnou 4 a 5.

Určujúcim faktorom pri rozhodovaní či je vzor spojitý alebo nie je vzdialenosť pixelov, ktoré ho tvoria. Pre každú dvojicu susedných pixelov spojitého vzoru platí, že ich vzdialenosť d nesmie byť väčšia ako $\sqrt{2}$ px.

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{4.3}$$



Obrázok 4–6: Nespojitý vzor.

4.2.1 Implementácia

Pre dosiahnutie rozlišovania spojitosti nájdených vzorov sme museli algoritmus Houghovej transformácie upraviť v dvoch miestach.

Prvú modifikáciu sme urobili už v prvom kroku algoritmu, kde vytvárame houghov priestor a prechádzame všetky pixely vstupnej matice. Pri vytváraní Houghovho priestoru si ukladáme nie len hodnotu daného pixelu ale aj jeho súradnice, ktoré ďalej využijeme pri výpočte vzdialenosti susedných pixelov. Vo výpise 4 sa nachádza telo funkcie, ktorá zo vstupnej matice 36 x 36 vytvára houghov priestor tvorený dvojrozmerným poľom spojkových zoznamov. Každý spojkový zoznam predstavuje množinu pixelov nachádzajúcich sa na jednej priamke, tak ako nasledujú za



Obrázok 4-7: Nespojitý vzor (pixely s hodnotu > 3).

sebou vo vstupnej matici. Indexami poľa sú parametre priamky, na ktorej sa dané body nachádzajú.

4. Výpis: Modifikovaná funkcia vytvárajúca Houghov priestor.

```
for (int x = 0; x < 36; x++) {
1
    for (int y = 0; y < 36; y++) {
2
      if (matrix[x][y] > 3) {
3
        for (int t = 0; t < ANGLES; t++) {
4
          double theta = (t + 1) * 15 * 3.14159 / 180;
5
          double r = x * cos(theta) + y * sin(theta);
6
          int r_index = round(r * 10) + 500;
7
          Pattern * pattern = (Pattern *)malloc(sizeof(Pattern));
8
          pattern \rightarrow x = x;
9
```

Druhú modifikáciu bolo potrebné urobiť pri prehľadávaní už vytvoreného houghovho priestoru. Pri prechádzaní spojkového zoznamu s pixelmi, porovnávame vzdialenosti dvoch po sebe nasledujúcich pixelov. Ak je ich vzdialenosť väčšia ako stanovená hranica je aktulný pixel posledným bodom aktálne nájdeného vzoru a nasledujúci pixel je počiatočným bodom ďalšieho nájedného vzoru.

5. Výpis: Modifikovaná funkcia prehľadávajúca houghov priestor.

```
for (int kq = 0; kq < ACC_SIZE; kq++) {
1
    for (int t = 0; t < ANGLES; t++) {
2
       Pattern *cur = accumulator[kq][t];
3
       int pattern_len = 0;
5
       int pattern_sum = 0;
6
       double pattern_avg = 0.0;
7
       int steps_control = 0;
8
       Pattern * prev = NULL;
9
       int dx = 0, dy = 0;
10
       bool koniec = false;
11
12
       int pattern_avg_idx = 0;
13
       while (cur != NULL) {
14
         steps_control++;
         pattern\_len++;
16
         pattern_sum += cur->value;
17
18
```

```
prev = cur \rightarrow prev;
19
20
            if (cur->prev != NULL) {
21
22
              dx = abs(cur \rightarrow x - prev \rightarrow x);
23
              dy = abs(cur \rightarrow y - prev \rightarrow y);
24
              dx = dx;
25
              dy = dy;
26
              if (\operatorname{sqrt}(\operatorname{dx} + \operatorname{dy}) > \operatorname{sqrt}(2)) {
27
                 koniec = true;
28
              }
29
            }
30
31
            if (cur->prev == NULL) koniec = true;
32
33
            if (koniec) {
34
              if (pattern_len > 1) {
35
                 pattern_avg = pattern_sum / (double) pattern_len;
36
                 pattern_avg_idx = round(pattern_avg * 10) - 30;
37
                 result_table [pattern_len - 2] [pattern_avg_idx]++;
38
              }
39
               pattern_len = 0;
40
              pattern_sum = 0;
41
              pattern_avg = 0.0;
42
               koniec = false;
43
            }
44
            \operatorname{cur} = \operatorname{prev};
45
         }
46
      }
47
48 }
```

4.2.2 Výsledky

Pre testovanie modifikovaného algoritmu houghovej transformácie boli použité dáta pochádzajúce zo simulácii UV pozadia pre vyhodnocovanie FTR. Pri testovaní spúšťačov sú vždy pri prekročení prahových hodôt vygenerované dva súbory. Jeden pre PTT spúšťač a druhý pre LTT spúšťač. Modifikovaný algoritmus sme testovali na dátach, ktoré prešli druhým stupňom spúšťačov (LTT) a je teda pravdepodobné, že sa v nich budú nachádzať vzory s väčšou dĺžkou.

Súbor s dátami z LTT spúšťača obsahuje snímky (matice) s rozmermi 36 x 36 px. Pre každý aktivovaný spúšťač sa do súboru uloží 31 matíc. Aktuálna snímka, 15 predchádzajucích a 15 nasledujúcich.



Obrázok 4–8: Histogram hodnôt pixelov z LTT súborov.

Na obrázku 4-8 môžeme vidieť histogram hodôt pixelov snímkov, nachádzajúcich sa v LTT súboroch. Z histogramu vidieť, že distribúcia hodnôt pixelov

sa riadí Poissonovým rozdelením. Na obrázkku 4-9 je histogram stredných hodnôt snímkov.



Obrázok 4–9: Histogram stredných hodnôt matíc z LTT súborov.

Závislosť počtu nájdených vzorov od ich priemernej hodnoty pre rôzne dĺžky je znázornená na obrázku 4-10. Ako môžeme vidieť, najdlhšie nájdené vzory mali dĺžku 6px a ich priemerná hodnota je 4,3. Vzory nájdené takýmto spôsobom nemôžu byť nikdy zamenené s reálnou spŕškou, keďže tá vytvára na snímkach pohybujúci sa bod. Keďže v dátach z LTT spúšťača je zahrnutý aj vplyv samotného detektora (vnesený šum) na zachytené UV spektrum, zistili sme touto simuláciou, že UV pozadie v kombinácii s vneseným šumom pochadzajúcim zosamotného detektora nevytvára vzory dlhšie ako 6px.



Obrázok 4–10: Závislosť počtu nájdených vzorov od dĺžky a priemernej hodnoty vzoru.

4.3 Houghova transformácia pre pohybujúce sa vzory

V našom prípade je v strede ohniskovej plochy detektora veľkosť pixelu na povrchu Zeme 0,51 km a pomaly rastie až ku vonkajšej hranici zorného poľa detektora (jeho krajným pixelom), kde dosahuje rozmer 0,61 km[23]. Ak predpokladáme spŕšku sekundárneho kozmického žiarenia šíriacu sa rýchlosťou svetla v atmosfére, potom počas 2,5.10⁻⁶ sekúnd dlhého GTU (Gate Time Unit) prejde spŕška dráhu dlhú 750 metrov. Čelo spŕšky sekundárneho kozmického žiarenia vytvorené primárnou časticou ultravysokej energie tvoria sekundárne častice s relativistickými energiami. Ako také sa šíria rýchlostou veľmi blízkou rýchlosti svetla a teda prijatý predpoklad o šírení sa častíc spŕšky rýchlostou svetla môžme pokladať za oprávnený. Obrázok 4-11 zachytáva situáciu pod deketorom pri prieniku spŕšky kozmického žiarenia atmosférou[24][25].



Obrázok 4–11: Spŕška zachytená JEM-EUSO teleskopom[21].

V súradnej sústave spojenej s detektorom potom vzor spŕšky bude zvierať uhol β s osou z detektora a uhol α s osou x detektora[26]. Situáciu popisuje obrázok 4–12. Červenou farbou je znázornená spŕška a zelenou jej priemet, ktorý uvidí detektor ako bod pochybujúci sa rýchlosťou blízkou rýchlosti svetla.

Na základe týchto poznatkov môžeme vytvoriť stratégiu skladania matíc (obrazcov) pre rozpoznávanie pohybujúceho sa vzoru. Základom stratégie je výpočet, ktorého výsledkom je pozícia (rozsvietený pixel) spŕšky v danom čase a pre jednotlivé uhly (α , β).



Obrázok 4–12: Geometria spŕšky.

$$x = c \cdot dt \cdot k \cdot \cos(\alpha) \cdot \cos(\beta)$$

$$c = 299\ 792\ 458\ m/s$$

$$dt = 2,5 \cdot 10^{-6}\ sec$$

$$k = GTU\ (cas)$$

$$(4.4)$$

Vzťahy 4.4 predstavujú závislosť pozície spŕšky od času (GTU) pre konkrétne uhly α a β .

4.3.1 Implementácia

Skladanie prebieha nasledovným spôsobom. Pre uvažovanú dvojicu uhlov α a β vypočítame zo vzorca 4.4 pozície spŕšky v jednotlivých časoch GTU. V našom prípade od 1 až po 31, pretože LTT spúšťač, ak je aktivovaný prepustí vždy 31 matíc (snímkov). Aktuálny snímok, 15 predchádzajucích a 15 nasledujúcich. Výsledkom výpočtu je tabuľka, ktorá údáva pozíciu spŕšky na osi x pre jednotlivé časy GTU. V našom prípade čas GTU predstavuje maticu a poloha na osi x stĺpec, ktorý sa má nachádzať vo výslednej matici.

Následne podľa vypočítaných pozícií vyskladáme výslednú maticu, pričom ak je medzi dvomi časmi rozdiel viac ako 1 pixel, chýbajúce stĺpce zoberieme z aktuálnej matice. Ak je výsledná matica menšia ako 36 x 36 px zvyšné stĺpce doplníme z

KPI

poslenej matice. Na obrázku 4-13znázornený proces skladania matíc aj s výslednou maticou. Pre jednoduchosť je na obrázku ukázaných len prvých 6 matíc.



 $\mathbf{Obrázok}\ \mathbf{4}-\mathbf{13}:$ Skladanie matic.

Pri vyskladávaní sme pre uhly α a β zvolili krok 15 stupňov, pričom sme prehľadávali jeden kvadrant, tak ako je to znázornené na obrázku 4–12. Pre jednu

FEI

žiarenia pohybovat atmosférou.

4.3.2 Výsledky

Ako vstupné dáta pre Houghovú transformáciu pre pohybujúce sa vzory boli použité LTT súbory vygenerované simuláciou UV pozadia, ktorá predstavovala 2500 sekúnd behu JEM-EUSO teleskopu. Vyskladané matice sme prehľadávali modifikovanou verziou houghovej transformácie. Na obrázku 4-14 môžeme vidieť závislosť počtu nájdených vzorov od dĺžky vzoru. Môžeme si všinúť, že vyskladávaním vznikli aj vzory dlhšie ako 6px. Na obrázku 4-15 je znázornená závislosť počtu nájdených vzorov od ich priemernej hodnoty pre jednotlivé dĺžky.



Obrázok 4–14: Závislosť počtu nájdených vzorov od dĺžky vzoru.

Na obrázkoch 4–17 až 4–22 je znázornená závislosť počtu nájdených vzorov od nastavenia uhlov α a β pre jednotlivé dĺžky nájdených vzorov. Môžeme si všimnúť,



Obrázok 4–15: Závislosť počtu nájdených vzorov od dĺžky a priemernej hodnoty vzoru.

že všetky získane grafy sú symetrické. Táto symetria vyplýva zo vztahu 4.4, kde pozícia spŕšky v čase t zavisí od súčinu kosínov uhlov α a β . To znamená, že bod

pohybujúci sa po ohniskovej ploche detektora predstavujúci spŕšku s nastavením uhlov $\alpha = 60$ a $\beta = 30$ prejde rovnakú trajektóriu ako bod predstavujúci spŕšku s nastavením uhlov $\alpha = 30$ a $\beta = 60$ a teda aj stratégia skladania je pre ne rovnaká. V dôsledku to zanamená, že v tomto štádiu je každý vzor, ktorý bol nájdený pre jednu kombináciu uhlov nájdený aj pre kombináciu opačnú.



Obrázok 4–16: Závislosť počtu nájdených vzorov s dĺžkou 2 p
x v závislosti od nastavenia uhlov α a β .



Obrázok 4–17: Závislosť počtu nájdených vzorov s dĺžkou 3 px v závislosti od nastavenia uhlov α a β .



Obrázok 4–18: Závislosť počtu nájdených vzorov s dĺžkou 4 p
x v závislosti od nastavenia uhlov α
a β .



Obrázok 4–19: Závislosť počtu nájdených vzorov s dĺžkou 5 p
x v závislosti od nastavenia uhlov α
a β .



Obrázok 4–20: Závislosť počtu nájdených vzorov s dĺžkou 6 p
x v závislosti od nastavenia uhlov α
a β .



Obrázok 4–21: Závislosť počtu nájdených vzorov s dĺžkou 7 px v závislosti od nastavenia uhlov α a β .



Obrázok 4–22: Závislosť počtu nájdených vzorov s dĺžkou 8 px v závislosti od nastavenia uhlov α a β .

Časový priebeh štyroch vybraných spŕšiek nájdených v testovacích dátach z LTT spúšťačov je znázornený na obrázku 4-23.V grafoch je na na osi x vynesený čas GTU (1 - 31) a na osi y suma hodnôt pixelov nad hodnotou pozadia (2) v aktuálnom bode vzoru. Za aktuálny bod sa berie aktuálny pixel vzoru a jeho okolie tvorené tromi najbližšími pixelmi. Tieto vybrané spŕšky boli tvorené vzorom s dĺžkou 8 px. Ako môžeme vidieť ich časový priebeh vyznačený modrou farbou nie je podobný s priebehom reálnej spŕšky, ktorá je v grafoch vyznačená fialovou farbou a teda nemôžu byť zamené s reálnou spŕškou.



Obrázok 4–23: Časový priebeh štyroch vybraných spŕšiek.

5 Záver

Finálnym výsledkom práce je odhad pravdepodobnosti zámeny falošného vzoru vytvoreného náhodne v obrovskom počte obrazcov so vzorom vytvoreným spŕškou kozmického žiarenia. Houghova metóda rozpoznávania vzorov hľadajúca vzor s charakteristikami spŕšok priniesla informáciu o počte nájdených vzorov s rôznou dĺžkou. Závislosť celkového počtu nájdených vzorov od ich dĺžky pre fyzikálne odvodený spôsob skladania matíc so simulovaným Poissonovským šumom je možné aproximovať štatisticky motivovanou funkciou:

$$N_p \,.\, (L_p) \sim (\frac{1}{N_{px}})^{L_p}$$
 (5.1)

kde N_p je počet nájdených vzorov, L_p je dĺžka vzoru a N_{px} je počet možných hodnôt jedného pixelu. Počet možných hodnôt pixelu, ako je vidieť z histogramu na obrázku 4-8 je 8, respektíve pre štatistiku z ktorej sa tento histogram vyhodnocoval je to 8 hodnôt. Pre väčšiu štatistiku sa môže vyskytnúť malý počet prípadov s jednou či viacerými hodnotami pixelu naviac. Ako základnú aproximáciu použijeme funkciu pre 8 možných hodnôt pixelu, ktorá zároveň odhaduje možnosť nájdenia falošného vzoru konzervatívnejšie ako funkcie pre výšší počet hodnôt.

Výsledok vyhodnocovania simulácie (modrá čiara so štvorcami) pozadia pre jeden PDM merajúci 2500 sekúnd (10^9 GTU) spolu s jej aproximáciou (červená čiara) je prezentovaný na obrázku 5–2. Vidíme, že pre štatistiku zodpovedajúcu približne trištvrťhodine meraní na jednom PDM môžeme nájsť desiatky vzorov s dĺžkou 8 pixelov.

Výsledok zo simulácie meraní (tmavomodrá čiara s trojuholníkmi) za 458 hodín (660 . 10^9 GTU) na jednom PDM, čo zodpovedá 3,3 hodinám meraní celého detektora vidíme na obrázky 5–3. Výsledok je aproximovaný dvoma funkciami. Prvá predpokladá 8 možných hodnôt pixelov, druhá deväť možných hodnôt. Pokiaľ aproximácia pre 8 možných hodnôt simulácie lepšie aproximuje časť krivky pre kratšie vzory, aproximácia pre 9 možných hodnôt pixelov časť krivky pre dlhšie vzory. Pre vyvodenie záverov o maximálnej možnej dĺžke vzoru pri dlhších meraní ostaneme pri aproximácii s 8 možnými hodnotami pixelov. Táto odhaduje výšši počet dlhších vzorov pre vyššie statistiky a tak nastavuje náš odhad konzervatívne, určuje spoľahľivejšie hornú hranicu odhadu dĺžky falošného vzoru.

Keďže štatistika použitá pre analýzu rozpoznávania vzorov zahŕňa obdobie 3.3 hodín meraní detektora, maximálnu dĺžku a počet vzorov pri dlhodobejších meraniach môžeme približne odhadnúť prenásobením aproximačnej funkcie faktorom (24 / 3, 3) pre hodnoty jednodenného merania a faktorom (24 / 3, 3). 365 . 3 pre merania za dobu 3 rokov. Hodnota (24 / 3, 3) škáluje získané výsledky na celodenné merania. Prenásobenie počtom dní a rokov potom preškáluje odhad na celú dobu merania pre celý detektor.

Odhad nám určuje výskyt jedného vzoru dlhého menej než 12 pixelov počas jednodenného merania. To znamená vzor dlhý 6 kilometrov. Pre trojročné merania sa v šume objaví falošný vzor dlhý maximálne 15 pixelov (7,5 kilometrov).

Prvý bod vzoru spŕšky sekundárneho kozmického žiarenia sledovateľný JEM-EUSO detektorom, teda prvý pixel detektora rozpoznateľný nad pozadím pre primárnu časticu s energiou 5 . 10^{19} eV s inklináciou (uhol 90-Beta stupňov v Houghovej metóde) 30 stupňov je viditeľný vo výške 13 kilometrov[27].

Predpokladajme pre najdlhšie nájdené falošné vzory s dĺžkou 7,5 kilometra (tj. priemet spŕšky na Zemskom povrchu dlhý 7,5 km) prvý detektorom viditeľný pixel vo výške 13 kilometrov a posledný bod na Zemskom povrchu. To znamená falošnú spŕšku s inklináciou 29,98 stupňa. Z geometrie spŕšky potom možno odhadnúť, že falošný vzor je možné čo sa týka jeho dĺžky zameniť počas trojročného merania so skutočnou UHECR udalosťou s inklináciou maximálne 30 stupňov. Pre jednodenné meranie ho je možné zameniť s UHECR udalosťou s inklináciou maximálne 24,7 stupňa.

Porovnaním s aktuálnou presnosťou rekonštrukcie UHECR udalostí v prog-

KPI

ramovom balíku ESAF na obrázku 5–1 vidíme, že pri použitom nastavení spúšťača za falošnú UHCER udalosť nemôže zameniť s reálnou UHECR udalosťou. Rekonštrukcia v rámci pre misiu zadefinovaných parametrov, kde presnosť rekonštrukcie smeru príchodu primárnej častice je vyžadovaná s odchýlkou menšou ako 2,5 stupňa, je možná pre častice tesne nad GZK limitom len pre primárne častice s inklináciou väčšou ako 30 stupňov.



Obrázok 5-1: Chyba gama-68 rekonštrukcie uhla UHECR udalostí v meraniach JEM-EUSO experimentu ako funkcia inklinácie primárnej častice voči povrchu atmosféry. Gama-68 je odchýlka 68 najbližších rekonštrukcií zo 100 simulovaných eventov. [1].

Ďalším výsledkom podporujúcim záver nemožnosti zameniť reálnu UHECR udalosť s falošnou je priebeh intenzity vzoru počas jeho šírenia sa atmosférou. Reálna spŕška má na prvých hodnotách svojho vzoru malé postupne do maxima rastúce hodnoty, ktoré následne klesajú. Ak sa pozrieme na priebeh vzorov nájdených našou analýzou tak tieto majú náhodný priebeh bez maxima v priebehu registrovaného vzoru. FEI



Obrázok 5 – 2: Závislosť počtu náj-
 Obrázok 5 – 3: Aproxmácia závis-

dených vzorov od ich dĺžky pre simu- losti počtu nájdených vzorov od ich

láciu 2500s meraní JEM-EUSO tele- dĺžky pre dlhšie obdobia meraní JEM-

skopu.skopu.EUSO teleskopu.

Zoznam použitej literatúry

- MERNIK, T.: Reconstruction of UHECR Events for the JEM-EUSO Mission: Diplomová práca. Tübingen: 2009. 129 s.
- [2] EBISUZAKI, Toshikazu et al.: Report on the Phase A Study 2010. Riken: JEM-EUSO Collaboration, 2010. 437 s. JEUSO-110025-01-E-TR-ZZZ
- [3] BLASCHKE, Filip: Analýza korelovaných spršek kosmického záření: Diplomová práca. Opava: SLU FPF, 2009. 103s.
- [4] BOBIK, Pavol: Oficiálna stránka Slovenskej účasti v JEM-EUSO experimente [online]. Košice: SAV, UEF, 2013. [s.a.] [cit. 2013-11-04]. Dostupné na internete: ">http://jem-euso.sk/JEMEUSO/>.
- [5] PASTIČÁK Blahoslav, BOBÍK Pavol, KUDELA Karel: JEM-EUSO experiment for extreme energy cosmic ray observation. In: Physics in Collision, Košice: SAV, 2013
- [6] JEM-EUSO Collaboration: JEM-EUSO : Extreme Universe Space Observatory... [online]. Riken, JEM-EUSO Collaboration, 2013. [s.a.] [cit. 2013-14-04].
 Dostupné na internete: ">http://jemeuso.riken.jp/en/.
- [7] KAJINO F. et al.: Overview of the JEM-EUSO Instruments. In: The JEM-EUSO Mission: Status and Prospects in 2011, Beijing, The JEM-EUSO Collaboration, 2011, s. 15 - 18. arXiv:1210.6797v3 [astro-ph.IM]
- [8] HACHISU Y. et al.: JEM-EUSO lens manufactoring In: The JEM-EUSO Mission: Status and Prospects in 2011, Beijing, The JEM-EUSO Collaboration, 2011, s. 83 - 86. arXiv:1210.6797v3 [astro-ph.IM]
- [9] KAWASAKI Y. et al.: The Focal Surface of the JEM-EUSO Instrument. In: The JEM-EUSO Mission: Status and Prospects in 2011, Beijing, The JEM-EUSO Collaboration, 2011, s. 31 - 34. arXiv:1210.6797v3 [astro-ph.IM]

- [10] THEODORIDIS, S., KOUTRUMBAS, K.: Pattern Recognition. Londýn: Elsevier Inc., 2009. 967 s. ISBN: 978-1-59749-272-0
- [11] POLIKAR, R.: Pattern Recognition In: Encyclopedia of Biomedical Engineering. Glassboro, New Jersey : Rowan University, 2006
- [12] DUDA R., Hart P., STORK D. Pattern Classification New York: Wiley, 2000,2. doplnené vydanie
- [13] JAIN A. K., DUIN R. P. W., MAO J.: Statistical pattern recognition: a review.In: IEEE Trans. Pattern Analysis Machine Intell., 2000
- [14] JAIN A., ZONGKER D.: Feature selection: evaluation, application, and small performance. In: IEEE Trans. Pattern Analysis Machine Intell., 1997
- [15] CARPENTER G. A., GROSSBERG S.: The ART of adaptive pattern recognition by a self-organizing neural network. In: Computer. roč. 21, 1998, s. 77
 - 88
- [16] CHALUPNIK V.: Biologické algoritmy (4) Neuronové sítě [online]. Praha: Internet Info, s.r.o., 2012. [s.a.]. [cit. 2013-16-04]. Dostupné na internete: <http://www.root.cz/clanky/biologicke-algoritmy-4-neuronove-site//>. ISSN 1212-8309
- [17] JEM-EUSO Collaboration: Expected performance of the JEM-EUSO mission: the exposure in nadir mode
- [18] BAYER J. et al.: The Cluster Control Board of the JEM-EUSO mission. In: The JEM-EUSO Mission: Status and Prospects in 2011, Beijing, The JEM-EUSO Collaboration, 2011, s. 51 - 54. arXiv:1210.6797v3 [astro-ph.IM]
- [19] FENU F.: A simulation study of space based mission for Ultra Hight Energy Cosmic ray search: Dizertačná práca. Kepler center-IAAT Tübingen: EKU, 2008. 149 s.

- [20] MERNIK T. et al.: Reconstruction of Extreme Energy Cosmic Ray Events Observed by JEM-EUSO in the ESAF Framework. In: Proceedings of the International Cosmic Ray Conference, Lodz 2009
- [21] MERNIK T. et al.: The ESAF Reconstruction Framework of UHECR Events for the JEM-EUSO Mission. In: The JEM-EUSO Mission: Status and Prospects in 2011, Beijing, The JEM-EUSO Collaboration, 2011, s. 119 - 122. arXiv:1210.6797v3 [astro-ph.IM]
- [22] DUDA O. R., HART E. R.: Use of the Hough Transformation To Detect Lines and Curves in Pictures. In: Communications of the ACM. roč. 15, 1972, č. 1, s. 11 - 15
- [23] TAKAHASHI Y. et al.: The JEM-EUSO mission. In: New Journal of Physics. roč. 11, 2009, č. 6
- [24] ABREU P. et al.: The Pierre Auger Observatory II: Studies of Cosmic Ray Composition and Hadronic Interaction models, Contributions to the 32nd International Cosmic Ray Conference, Beijing, China, August 2011, ar-Xiv:1107.4804 [astro-ph.IM]
- [25] GRIEDER P.K.F.: Extensive Air Showers. Springer-Verlag, Berlin Heidelberg 2010
- [26] ADAMS J. H. Jr et al.: The JEM-EUSO Mission: Status and Prospects in 2011. In: The JEM-EUSO Mission: Status and Prospects in 2011, Beijing, The JEM-EUSO Collaboration, 2011. arXiv:1204.5065 [astro-ph.IM]
- [27] MERNIK T.: osobná komunikácia.
- [28] MUSIL C., TRIPIŠOVÁ B.: MERANIA NA OPTICKÝCH SÚSTAVÁCH [online]. Žilina: UNIZA, EF, 2012. [s.a.]. [cit. 2013-03-05]. Dostupné na internete: http://fyzika.uniza.sk/ trpisova/skriptaopticke_sustavy.pdf>.
- [29] JARKOVSKÝ J., LITTNEROVÁ S.: Vícerozměrné statis-

[online]. tické metody: Diskriminační analýza Brno: MUNI, 2013-03-05]. IBA, 2011.[s.a.]. [cit. Dostupné na internete: <http://www.iba.muni.cz/esf/res/file/bimat-prednasky/</pre> vicerozmerne-statisticke-metody/VSM-08.pdf>.

[30] DAŇO I., OSTERTAGOVÁ E.: Numerické metódy, pravdepodobnosť a matematická štatistika. Prvé vydanie. Košice: EQUILIBRIA, 2010. 166 s. ISBN 978-80-89284-56-6

Zoznam príloh

- **Príloha A** Používateľská príručka
- **Príloha B** Systémová príručka
- Príloha C CD médium

Technická univerzita v Košiciach Fakulta elektrotechniky a informatiky

Algoritmy pre rozpoznávanie obrazcov pri štúdiu kozmického žiarenia ultravysokých energií v rámci JEM-EUSO experimentu

Príloha A

Používateľská príručka

 $\mathbf{2013}$

Bc. Miroslav Staroň

Program pre analýzu LTT súborov

Funkcia programu

Hlavnou úlohou programu je vyhľadávanie vzorov (spŕšiek) v súboroch vygenerovaných LTT spúšťačom pomocou Houghovej transformácie a vyskladávania matíc.

Zoznam priložených súborov na CD

Aplikácia sa nachádza na priloženom CD, ktoré obsahuje zdrojové kódy aplikácie, spustiteľné súbory ako aj používateľskú príručku vo formáte PDF.

Zoznam priložených súborov na CD:

- dist/hough.cpp-zdrojový kód programu, v ktorom je implementovaná Houghová transformácia
- doc dokumentácia

Minimálne HW požiadavky

- CPU Pentium 4
- 2048 MB RAM
- 2GB voľného miesta na disku
- CD/DVD mechanika

SW požiadavky

- operačný systém Linux alebo Unix
- C++ kompilátor
Inštalácia programu

Program je potrebné pred použitím skompilovať. Napríklad v OS Linux príkazom:

g++ -O2 -march=native -o hough hough.cpp

Používanie programu

Program spustíme príkazom ./hough [zoznam parametrov].

- -h, --help zobrazí nápovedu
- -i, --input <LTT subor> cesta k LTT súboru, ktorý bude analyzovaný (povinný parameter)
- -o, --output <výstupný subor> cesta k súboru, do ktorého bude zapísana štatistika nájdených vzorov (povinný parameter)
- -t, --treshold <dĺžka vzoru> parameter určuje dlžku vzorov pre ktorú bude pre každý nájdený vzor zapísaná podrobnejšia štatitika do samostaného priečinku
- -p, --prefix <predpona priečinku> prepdona použitá pre názvy priečinkov s podrobnou štatistikou nájdených vzorov

Formát súborov

Vstupný LTT súbor obsahuje štyri stĺpce, pričom v prvom je x-ová súradnica, v druhom y-ová súradnica, v treťom čas (GTU 0 - 30) a v poslednom hodnota pixelu. Pre každý uložený spúšťač sa v súbore nachádza 40176 riadkov.

- 1 0 0 3
- 1 0 1 1
- 1 0 2 1

1 0 3 4

1 0 4 2

Výstupný súbor so štatistikou s nájdenými vzormi obsahuje päť stĺpcov. V prvom je uložený uhol α , v druhom uhol β , v treťom dĺžka vzorov, vo štvrtom priemerná hodnota vzoru a v poslednom je počet nájdených vzorov s takýmito parametrami.

0	1	1	9	0
0	1	1	10	265
0	1	1	11	0
0	1	1	12	0
0	1	1	13	323

Priečinok s podrobnou štatistikou obsahuje štyry textové súbory:

- info.txt s informáciami o uhloch, dĺžke a priemernej hodnote nájdeného vzoru
- coordinates.txt obsahuje súradnice pixelov nájdeného vzoru
- LTT_data.txt všetkých 31 snímkov(matíc 36 x 36 px) ktoré prepustil LTT spúšťač (40176 riadkov)
- development.txt časový priebeh spŕšky

Program pre zobrazenie výsledkov z analýzy LTT súborov

Funkcia programu

Hlavnou úlohou programu je zobrazenie grafov závislosti počtu nájdených vzorov od nastavenia uhlov α a β pre jednotlivé dĺžky nájdených vzorov (2 - 8).

Zoznam priložených súborov na CD

Aplikácia sa nachádza na priloženom CD, ktoré obsahuje zdrojové kódy aplikácie a používateľskú príručku vo formáte PDF.

Zoznam priložených súborov na CD:

- dist/viewer.cpp zdrojový kód programu pre zobrazenie výsledkov
- doc dokumentácia

Minimálne HW požiadavky

- CPU Pentium 4
- 1024 MB RAM
- 50MB voľného miesta na disku
- CD/DVD mechanika

SW požiadavky

- operačný systém Linux alebo Unix
- \bullet C++ kompilátor
- nainštalovaný root framewrok <http://root.cern.ch/>

Inštalácia programu

Program je potrebné pred použitím skompilovať. Napríklad v OS Linux príkazom:

g++ -O2 -march=native -o viewer viewer.cpp 'root-config --cflags --libs'

Používanie programu

Program spustíme príkazom ./viewer <cesta k priečinku s výsledkami>.

Obrázok 5–1: Spustený program pre zobrazene výsledkov analýzy.

Technická univerzita v Košiciach Fakulta elektrotechniky a informatiky

Algoritmy pre rozpoznávanie obrazcov pri štúdiu kozmického žiarenia ultravysokých energií v rámci JEM-EUSO experimentu

Príloha B

Systémová príručka

 $\mathbf{2013}$

Bc. Miroslav Staroň

Štruktúra programu

Program je implementovaný ako konzolová aplikácia v jazyku C++.

Použité dátové štruktúry

Štruktúra **Bunch** je použitá ako spojkový zoznam, do ktorého sa po spustení programu načíta celý vstupný súbor.

```
struct Bunch {
    int matrix[31][36][36];
    bunch * next;
};
```

Štruktúra Point slúži ako spojkový zoznam bodov, ktoré majú zhodné parametre (r, θ) a ležia teda na jednej priamke.

```
struct Point {
    int x;
    int y;
    int value;
    Pattern * prev;
};
```

Štruktúra **Eas** slúži na uchovanie informácií o nájdenej spŕške kozmického žiarenia.

```
struct Eas {
    int pattern_length;
    double pattern_average;
```

```
int * x_coords;
int * y_coords;
int * values;
int alpha_idx;
int beta_idx;
```

};

Zoznam funkcií

calcPositions

Funkcia calcPositions vypočíta na základe parametrov (α , β) pozíciu spŕšky pre jednotlivé časy GTU (1 - 31). Výsledok uchová v dvoch globálnych premenných. Prvá int columns[7][7][36] slúži na uchovanie indexu stĺpca a druhá int matrices[7][7][36] na uchovanie indexu matice, z ktorej má byť stĺpec vybratý. Indexmi v týchto poliach sú uhly α , β a index stĺpca vo výslednej matici.

void calcPositions(int alpha, int beta)

get Matrix Bunch From File

Funkcia getMatrixBunchFromFile načíta 40176 riadkov zo vstupného súboru špecifikovanom v parametri infile. Tento počet riadkov predstavuje jednu uloženú udalosť, pre ktorú bol spustený LTT spúšťač. Načítané riadky sú uložené do poľa matrixBunch, ktoré je parametrom funkcie. Celý vstupný súbor je načítaný postupným volaním tejto funkcie.

```
bool getMatrixBunchFromFile(ifstream &infile,
```

int matrixBunch[31][36][36])

buildMatrix

Funkcia buildMatrix na základe parametrov (uhlov α a β) vyskladá novú maticu. Súbor pôvodných matíc je funkcii odovzdaný v parametri matrixBunch a nová výsledná matica je uložená do poľa špecifikovanom v parametri matrix.

```
void buildMatrix(int matrix[36][36],
```

int matrixBunch[31][36][36],
int alpha,
int beta)

buildHoughSpace

Funkcia buildHoughSpace vypočíta Houghov priestor z matice matrix s rozmermi 36 x 36 px. Vytvorený Houghov priestor je uložený do poľa accumulator.

```
void buildHoughSpace(int matrix[36][36],
```

Point * accumulator[ACCSIZE][ANGLES])

searchInHoughSpace

Funkcia searchInHoughSpace prehľadáva skôr vytvorený Houghov priestor odovzdaný funckii refernciou v parametri accumulator. Výsledky sa ukladajú do globálneho poľa resultTableAngles.

saveEvent

Funkcia saveEvent uloží informácie o nájdenom vzore do priečinku s menom <prefix>.<ID>. Prefix názvu priečinku je špecifikovaný parametrom pri spúšťaní programu. Informácie o nájdenom vzore sú funkcii odovzdané v parametri eas. Parameter matrixBunch je referenciou na pole pôvodných matíc. Vytvorený priečinok bude obsahovať nasledovné súbory:

- info.txt s informáciami o uhloch, dĺžke a priemernej hodnote nájdeného vzoru
- coordinates.txt obsahuje súradnice pixelov nájdeného vzoru
- LTT_data.txt všetkých 31 snímkov(matíc 36 x 36 px) ktoré prepustil LTT spúšťač (40176 riadkov)
- development.txt časový priebeh spŕšky

void saveEvent(struct Eas * eas, int matrixBunch[31][36][36])

processJob

Vo funkcii processJob prebieha prehľadanie jednej udalosti zo vstupného súboru. Pre každú zo 49 kombinácií uhlov α a β je vyskladaná výsledná matica, ktorá je odovzdaná funkcii buildHoughSpace a následne funkcii searchInHoughSpace.

void processJob(int matrixBunch[31][36][36])

printUsage

Funkcia printUsage vypíše nápovedu pre program.

void printUsage(ostream& stream, int exitCode)



Obrázok 5–1: Vývojový diagram funkcie processJob.

main

Funkcia main je hlavnou funkciou programu a je zavolaná hneď po spustení programu. Následne sa vstupný súbor načíta do globálnej premennej jobQueue a pre každú načítanú udalosť je zavolaná funkcia processJob.

Program pre zobrazenie výsledkov z analýzy LTT súborov

Štruktúra programu

Program je implementovaný ako grafická (GUI) aplikácia, pričom využíva API knižnic z aplikačného rámca CERN root. Program je implementovaný v jazyku C++.

5



 $\mathbf{Obrázok}\ \mathbf{5-2}:$ Vývojový diagram funkcie main programu hough .

Zoznam funkcií

${\bf readResults} From File$

Funkcia readResultsFromFile prečíta súbor vygenerovaný programom hough a prečítané záznamy uloží do globálneho poľa resultTableAngles.

void readResultsFromFile(ifstream &infile)

main

Funkcia main je hlavnou funkciou programu a je zavolaná hneď po spustení programu. Následne sa načítajú vstupné subory do globálnej premennej resultTableAngles a pre každú dĺžku vzoru je vykreslený graf so závislosťou počtu nájdených vzorov od uhlov α a β .

int main(int argc,

char ** argv)