

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

2014

Bc. Erik Gajdoš

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 2014

Bc. Erik Gajdoš

Abstrakt v SJ

JEM-EUSO je prvým experimentom využívajúci zemskú atmosféru ako obrovský detektor s kapacitou dostatočnou na vyriešenie otázky pôvodu a určenia charakteristík UHECR. Táto práca analyzuje možnosť výskytu falošnej spršky v simulovaných JEM-EUSO dátach, reprezentujúcich šum s Poissonovským rozdelením. Cieľom práce je nájsť metódu rozpoznávania vzorov vhodnú pre detekciu falošných spršok. Na základe výsledkov detekcie falošných spršok zistiť, či je možné si zameniť reálnu spršku s falošnou a integrovať metódu do kolaboračného balíčka ESAF.

Kľúčové slová

JEM-EUSO, UHECR, rozpoznávanie vzorov, zhlukovanie

Abstrakt v AJ

JEM-EUSO is the first experiment which uses Earth's atmosphere as a giant detector with capacity sufficient to resolve the question of UHECR origin and characteristics. The thesis analyse possibility of fake shower occurrence in poisson distributed noise of simulated JEM-EUSO background data. The goal of thesis is to find pattern recognition method for fake shower detection. Basis on the shower detection results determine, if is possible to mistake fake shower with real one and integrate method to collaboration framework ESAF.

Kľúčové slová v AJ

JEM-EUSO, UHECR, clustering, pattern recognition

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**
Algorithms for pattern recognition in the study of ultra-high energy cosmic rays
within the JEM-EUSO experiment

Študent: **Bc. Erik Gajdoš**
Školiteľ: **doc. Ing. Ján Genči, PhD.**
Školiace pracovisko: **Katedra počítačov a informatiky**
Konzultant práce: **RNDr. Pavol Bobík, PhD.**
Pracovisko konzultanta: **Ústav experimentálnej fyziky SAV**

Pokyny na vypracovanie diplomovej práce:

1. Oboznámiť sa so základmi fyzikálnej problematiky detekcie častíc s ultravysokými energiami.
2. Preštudovať si princípy fungovania JEM-EUSO experimentu so zameraním sa na rozlíšenie vzorov s charakteristikami spršok kozmického žiarenia v registrovanom UV signáli.
3. Aplikácia metódy zhlukovania na rozpoznanie falošných vzorov v uniformnom šume.
4. Aplikácia metódy zhlukovania na rozpoznanie vzoru pohyblivej spršky v uniformnom šume.
5. Implementovanie metódy zhlukovania na simulované JEM-EUSO dáta.
6. Porovnanie s výsledkami získanými Haughovou transformáciou-metódou.
7. Integrovanie metódy do kolaboračného frameworku ESAF.
8. Spracovať dokumentáciu podľa pokynov vedúceho práce.

Jazyk, v ktorom sa práca vypracuje: slovenský
Termín pre odovzdanie práce: 02.05.2014
Dátum zadania diplomovej práce: 31.10.2013


.....
doc. Ing. Jaroslav Porubán, PhD.
vedúci garantujúceho pracoviska




.....
prof. Ing. Liberios Vokorokos, PhD.
dekan fakulty

Čestné vyhlásenie

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

Košice 2. 5. 2014

.....

Vlastnoručný podpis

Pod'akovanie

Ako prvé, by som sa chcel poďakovať priateľom a rodine za ich podporu pri písaní tejto práce. Najviac sa chcem poďakovať konzultantovi diplomovej práce RNDr. Pavlovi Bobíkovi, PhD. za odborné vedenie, konzultácie a usmerňovanie pri vytváraní diplomovej práce. Taktiež sa chcem poďakovať aj vedúcemu diplomovej práce doc. Ing. Jánovi Genčimu, PhD. za odborné vedenie a cenné pripomienky.

Obsah

Úvod	1
1 Formulácia úlohy	3
2 Kozmické žiarenie	4
2.1 Energetické spektrum	5
2.2 GZK limit	6
2.3 Pôvod kozmického žiarenia	7
2.4 Spršky kozmického žiarenia	8
2.5 Pozorovateľné efekty	8
3 Sledovanie vysokoenergetického kozmického žiarenia	9
3.1 Experiment JEM-EUSO	9
3.2 Nástroje JEM-EUSO experimentu	10
3.3 JEM-EUSO a rozpoznávanie obrazcov	12
4 Zachytenie spršky JEM-EUSO detektorom	13
4.1 Prvá spúšťacia úroveň	14
4.2 Druhá spúšťacia úroveň	15
4.3 Udalosť určená na spracovanie	15
5 Rozpoznávanie vzorov	17
5.1 Klasifikácia vzorov	17
5.2 Adaptácia a učenie sa	19
5.3 Vyhodnotenie výsledkov	20
5.4 Zhluková analýza	20
5.5 PWISE metóda na vyhľadávanie spršok	21
5.6 Houghova transformácia pre všeobecné použitie	23

6	Kolaboračný rámec projektu ESAF	25
6.1	Simulačný balíček rámca ESAF	26
6.2	Rekonštrukčný balíček rámca ESAF	27
6.2.1	Moduly v rekonštrukčnom balíku rámca ESAF	28
6.2.2	Reprezentácia spřšky v softvérovom rámci ESAF	29
6.3	Softvérový balík ROOT	30
7	Metódy skladania udalosti v 2D	31
7.1	Návrh algoritmu pre vyskladanie matice	31
7.1.1	Vyskladanie matice jednoduchou sumáciou	32
7.1.2	Vyskladanie matice upravenou sumačnou metódou	35
7.2	Návrh algoritmu pre rozpoznávanie vzoru nad vyskladanou maticou	38
7.2.1	Návrh vzoru spřšky pre rozpoznávanie	38
7.2.2	Návrh spôsobu prechádzania vyskladanou maticou	39
7.3	Návrh algoritmu pre integráciu do ESAF rámca	40
7.4	Návrh pre implementáciu Houghovho priestoru	42
8	Implementácia algoritmov na rozpoznávanie vzoru	43
8.1	Implementácia rozpoznávania pomocou metódy zhlukovania	43
8.2	Implementácia správnosti použitia metódy zhluku	44
8.3	Testovanie pomocou metódy zhlukovania	46
8.3.1	Jediný vzor v jednom stłpci	47
8.3.2	Viacero vzorov v jednom stłpci	47
8.4	Implementácia modulu do rámca ESAF	48
8.5	Implementácia Houghovho priestoru a nespojitost' vzoru	49
8.6	Implementácia Houghovho priestoru a spojitost' vzoru	52
9	Overovanie výsledkov	54
9.1	Overenie vplyvu závislosti hranice šumu na výsledky	54

9.2	Vyhľadávanie spířšok v Poissonovom řume	56
9.3	Porovnanie výsledkov s vyskladávaním Houghovou transformáciou . .	57
10	Záver	59
	Zoznam použitej literatúry	62
	Zoznam príloh	66

Zoznam obrázkov

2-1 Spracovanie merania rôzneho spektra energií kozmického žiarenia. Bodkovaná čiara reprezentuje zákon mocniny E^{-3} . Taktiež sú uvedené integrálne toky[7].	6
3-1 Ilustračný pohľad na JEM-EUSO teleskop namontovaný na ISS.[14]	10
3-2 Frešnelové šošovky..[16]	11
3-3 Náčrt fungovania JEM-EUSO experimentu.[17]	12
4-1 Ukážka distribúcie fotoelektrónov na ohniskovej vzdialenosti povrchu pre spŕšku s energiou 10^{20} eV a zenitový uhol $\theta = 60^\circ$. [33]	13
5-1 Znázornenie šumu pozadia s maximálnou hodnotou 5 zmiešaného s hodnotami signálu, získaných metódou PWISE.[27]	21
5-2 Ukážka priamej línie (červená čiara) a jej kolmej vzdialenosti k počiatku súradnicovej sústavy (modrá čiara).[28]	23
6-1 Zobrazenie základného toku rámca ESAF.[33]	25
6-2 Zobrazenie štruktúry rekonštrukčného modelu rámca ESAF.[33]	28
7-1 Diagram postupov pre jednoduchý sumačný algoritmus.	33
7-2 Zobrazenie falošnej spŕšky, ktorú vložíme do udalosti pre otestovanie algoritmu.	34
7-3 Ilustračné zobrazenie vloženej falošnej spŕšky do udalosti a zosumarizovanej matice podľa jednoduchej sumácie.	35
7-4 Diagram postupov pre modifikovaný sumačný algoritmus.	36
7-5 Ilustračné zobrazenie vloženej falošnej spŕšky do udalosti a zosumarizovanej matice podľa modifikovanej sumácie.	37
7-6 Trojuholníky reprezentujúce kvadranty, ktoré sa použijú pri výpočte nasledujúceho bodu na základe uhla spŕšky.	40
8-1 Zobrazenie histogramu priemernej hodnoty pixelu matice.	45
8-2 Zobrazenie závislosti počtu nájdených spŕšok od priemernej hodnoty patternu a jeho dĺžky (5).	46

8-3	Zobrazenie teoretickej pravdepodobnosti počtosti a reálneho výsledku.	47
8-4	Zobrazenie teoretickej pravdepodobnosti počtosti a reálneho výsledku.	48
8-5	Najdlhšie nespojité vzory pre $M=72$ (ľavý panel obrázku) a $M=12$ (pravý panel).	50
8-6	Závislosť počtu nespojitých vzorov od dĺžky vzoru pre rôzne počty smerov v Houghovej metóde (viď. popis v texte) pre signál po pre- chode spúšťačou schémou.	50
8-7	Závislosť počtu nespojitých vzorov od dĺžky vzoru pre rôzne počty smerov v Houghovej metóde (viď. popis v texte) pre Poissonovský signál bez prechodu spúšťačom.	51
8-8	Porovnanie počtu vzorov pre udalosti vyselektované spúšťačom a ge- nerovaný signál(ľavý panel) a ich pomer.	52
8-9	Najdlhšie spojité vzory pre $M=72$ (ľavý panel obrázku) a $M=12$ (pravý panel).	53
9-1	Závislosť počtu vzorov od ich dĺžky pre rôzne hranice šumu.	55
9-2	Závislosť počtu vzorov od ich dĺžky získaných Houghovou metódou a metódou zhlukovania.	56
9-3	Závislosť počtu vzorov od ich dĺžky získaných z merania a vyskladávania pomocou Houghovej transformácie [35].	58

Zoznam tabuliek

4-1	Jednotlivé stupně spúšťáča a ich prislúchajúce časti [32].	14
-----	--	----

Zoznam symbolov a skratiek

AGASA Akeno Pole obrovskej spršky

CERN Európska organizácia pre jadrový výskum

EAS Extenzívna časticová sprška

ESA Európska vesmírna agentúra

ESAF Rámec JEM-EUSO kolaborácie pre simuláciu a analýzu

GTU Základná časová jednotka

GZK Greisen Zatsepin Kuzmin limit

HiRes Detektor Flyove oko kozmického žiarenia s vysokým rozlíšením

IDL Interaktívny dátový jazyk

ISS Medzinárodná vesmírna stanica

JAXA Japonská vesmírna agentúra

JEM-EUSO Kozmické Observatórium Extrémneho Vesmíru na palube Japonského
Experimentálneho Modulu

LHC Veľký hadrónový urýchľovač

LTT spúšťač lineárnej dráhy

NASA Národný úrad pre letectvo a kozmonautiku

PDM Detekčný modul fotónov

PDM Modul Svetelného Detektora

PWISE Technika pre vyhľadávanie okien a maxima

ROOT Rýchle Objektovo-Orientované Technológie

Roskosmos Ruská letecká a vesmírna agentúra

SNR Séria odstupu signálu od šumu

UHECR Kozmické žiarenie s ultravysokými energiami

UHECR Ultra High Energy Cosmic Ray

Slovník termínov

Mpc Megaparsek - Je astronomická jednotka vzdialenosti o veľkosti 1 000 000 parsek. Parsek je rovný približne 3.2616 svetelným rokom.

ISS Medzinárodná vesmírna stanica - Jediná trvalo obívaná stanica vo vesmíre. Nachádza sa na obežnej dráhe okolo Zeme vo výške okolo 410 kilometrov.

HTV H-II Prepravná loď - Japonská bezpilotná kozmická loď používaná na dopravu nákladu k medzinárodnej vesmírnej stanici.

UHECR Kozmické žiarenie s vysokou energiou - Kozmické žiarenie, ktorého kinetická energia je vyššia ako 10^{19} eV, čo je vysoko nad energiami bežnými v kozmickom žiarení.

ESAF EUSO Simulačný a analyzačný softvérový rámec - Softvérový rámec vyvíjaný kolaboráciou JEM-EUSO pre simuláciu analýzy a rozpoznávania vysokoenergetických kozmických žiarení v rámci JEM-EUSO experimentu.

Astronómia UHECR častíc Trajektórie častíc s ultravysokou energiou sú len minimálne ovplyvňované medzigalaktickými a galaktickými magnetickými poliami. Smer ich príchodu do detektoru je preto tamer rovnaký ako smer k ich zdroju. Vďaka tomu, je možné rovnako ako pri astronomickom pozorovaní fotónov využiť UHECR častice na vytvorenie mapy ich zdrojov. Na základe tejto analógie hovoríme o astronómii UHECR častíc.

CINT CINT je interpretérom jazyka C++ používanom v softvérovom balíku ROOT a je napísaný v jazyku C++. CINT je užitočný v prípadoch, keď rýchly vývoj aplikácie je podstatnejší ako čas vykonávania programu.

Úvod

Viacere problémy súčasnej astrofyziky vyžadujú riešenie problému rozpoznania vzorov vo veľkom množstve registrovaných údajov. Príkladom takejto úlohy je rozpoznanie signálu tvoreného spľškou kozmického žiarenia v snímkach fotocitlivých detektorov. V takýchto prípadoch sa hľadajú metódy rozpoznávania signálu využiteľné pri rôznych špecifických zariadeniach. Použitie rozličných metód závisí od dostupných zdrojov a vyžadovanej presnosti rozpoznávania.

JEM-EUSO je fyzikálny experiment zameraný na vyriešenie otázky pôvodu kozmického žiarenia ultravysokých energií (ďalej UHECR). UHECR je kozmické žiarenie s energiami nad tzv. Greisen Zatsepin Kuzmin (ďalej GZK) limitom ($>6.10^{19}$ eV) efektívne interagujúce s mikrovlnným pozadím vesmíru. UHECR častica vďaka tejto interakcii stráca energiu a na vzdialenostiach v rádoch desiatok megaparsekov ju zníži pod GZK limit. Registrácia UHECR častíc na Zemi vedie k záveru, že zdroje tohto žiarenia nie sú od Zeme vzdialené viac ako pár desiatok megaparsekov. Napriek päťdesiatim rokom výskumu tohto fenoménu, zdroje UHECR a mechanizmy jeho urýchľovania nie sú známe.

Medzi možné zdroje UHECR častíc patria pulzary, rádiogalaxie, supernovy a zdroje zábleskov gama žiarenia, alebo to môžu byť doposiaľ neznáme objekty. V blízkosti našej galaxie, vo vzdialenosti do 50 megaparsekov, sú to napríklad najjasnejšie rádiové galaxie Kentaur-A a Panna M-87. Ak merania v nasledujúcich rokoch neidentifikujú ako zdroj žiaden známy objekt, môže to vyvolať pochybnosti o platnosti špeciálnej teórie relativity alebo iných fundamentálnych fyzikálnych princípov.

JEM-EUSO experiment je spoločným projektom kozmických agentúr JAXA, ESA, NASA, Roskosmos a spolupracujúcich krajín. Celkový počet zapojených krajín je 13, vrátane Slovenskej republiky. Kľúčovým účastníkom JEM-EUSO projektu je Japonsko, ktoré poskytuje pre projekt základnú infraštruktúru zahŕňajúcu nosnú raketu HII-B, kozmickú loď HTV a pozíciu pre umiestnenie detektora na Medzinárodnej

kozmickej stanici ISS na Japonskom module Kibo. Predpokladaný dátum štartu je január 2017, ktorý bude nasledovaný minimálne trojročným meraním.

JEM-EUSO je prvý experiment využívajúci zemskú atmosféru ako obrovský detektor s kapacitou dostatočnou na vyriešenie otázky pôvodu a určenia charakteristík UHECR. Dosiahnutie výsledkov v rámci JEM-EUSO experimentu môže priniesť principiálne nové fyzikálne poznatky. Jednou z otázok, ktorú je potrebné vyriešiť je zistenie, či je pri analýze možné zameniť si reálnu spršku s falošnou.

JEM-EUSO experiment je vo fáze prípravy simulačného softvéru a aktuálne sa pracuje na metódach rozpoznávania spršok a analýzy pravdepodobnosti výskytu falošného signálu v simulovaných JEM-EUSO dátach. Táto úloha je konkrétne zodpovednosťou slovenskej skupiny JEM-EUSO kolaborácie. Riešenie tejto úlohy prebieha v spolupráci s Ústavom Experimentálnej Fyziky SAV. Pre toto zadanie je cieľom nájsť metódu detegujúcu spršku. Je vhodné preveriť viacero metód pre verifikáciu výsledkov a porovnanie presnosti detekcie falošných spršok.

1 Formulácia úlohy

JEM-EUSO experiment sa zameriava na sledovanie výskytu UHECR spršok v atmosfére Zeme. Pre návrh vhodnej metódy rozpoznávania je potrebné oboznámiť sa so základmi fyzikálnej problematiky detekcie častíc s ultravysokými energiami. Následne je cieľom pochopiť problematiku UHECR a oboznámiť sa s princípmi fungovania JEM-EUSO experimentu so zameraním sa na rozlíšenie vzorov s charakteristikami spršok kozmického žiarenia v registrovanom UV signále.

Po dôkladnom preštudovaní a oboznámení sa s fyzikálnou podstatou problému sa ako prvá pre overenie funkčnosti metódy aplikuje metóda zhlukovania na rozpoznanie falošných vzorov pre analyticky overiteľný prípad uniformného šumu. Následne po aplikácii rozpoznávania nad statickými dátami sa metóda zhlukovania aplikuje na rozpoznávanie vzoru pohyblivej spršky v uniformnom šume. Finálne sa metóda zhlukovania aplikuje na simulované JEM-EUSO dáta. Získané výsledky sa následne porovnajú s výsledkami získanými pomocou Houghovej transformácie-metódy. Po overení funkčnosti metódy zhlukovania sa táto metóda integruje do kolaboračného rámca ESAF.

Diplomová práca pristupuje k analýze simulovaných JEM-EUSO dát inkrementálne. Najskôr sa metódy rozpoznávania vzorov aplikujú na signál, v ktorom sa analyticky určí výsledok, čím sa overí funkčnosť použitej metódy. Následne sa metódy rozpoznávania iteračne blížajú k analýze reálnych dát. Uvedený prístup je zvolený so zámerom zabezpečenia kontroly správnosti analýzy.

2 Kozmické žiarenie

Kozmické žiarenie tvoria elementárne častice dopadajúce do zemskej atmosféry z vesmíru. Pomenovanie kozmické žiarenie má historický pôvod, v skutočnosti ide o korpuskulárne častice. Väčšinou ide o protóny a alfa častice, no prítomné sú aj jadrá ťažkých prvkov, elektróny a pozitrony. Na základe rôznych pozorovaní je jasné, že tieto častice dopadajú na zemský povrch zo všetkých strán rovnako, nezávisle na čase [1].

Posledné tvrdenie platí len pre častice s dostatočne veľkou energiou, keďže nabitá častica galaktického kozmického žiarenia predtým, ako bude detegovaná na Zemi, musí preletieť cez medziplanetárne magnetické pole a magnetické pole Zeme. Z Lorentzovej rovnice popisujúcej pohyb nabitej častice v magnetickom poli vidíme, že so vzrastajúcou energiou je dráha častice menej zakrivená. Preto častice s väčšou energiou (nad 100 GeV) preletia cez Slnecnú sústavu takmer bez zakrivenia ich trajektórie, častice s malou energiou sa špirálovito pohybujú pozdĺž siločiar magnetického poľa a niektoré sú magnetickým poľom Zeme zachytené[1].

Pred dopadom do atmosféry nazývame častice kozmického žiarenia primárnymi časticami. Pri ich zrážkach s molekulami atmosféry dochádza k vzniku veľkého množstva sekundárnych častíc. Pri zrážke primárnej častice s jadrom atómu molekuly vzduchu (napríklad s dusíkom ^{14}N) vznikajú viaceré typy sekundárnych častíc. Najčastejšie sú to nestabilné pióny, ktoré sa v krátkom čase rozpadajú na mióny. Tieto sa opäť zrážajú s molekulami atmosféry a produkujú ďalšie sekundárne častice. Množstvo vytvorených sekundárnych častíc závisí od energie primárnej častice a od jej typu. Tvorba sekundárnych častíc má lavínový charakter pri časticách s dostatočne veľkou energiou, čo označujeme ako časticová sprška (air shower). Ak má primárna častica dostatočne vysokú energiu, vtedy môže vzniknutá sprška dosiahnuť povrch Zeme, čo sa v literatúre označuje ako extenzívna časticová sprška (Extensive Air Shower)[2]. Metódy moderného výskumu UHECR sa opierajú najmä o

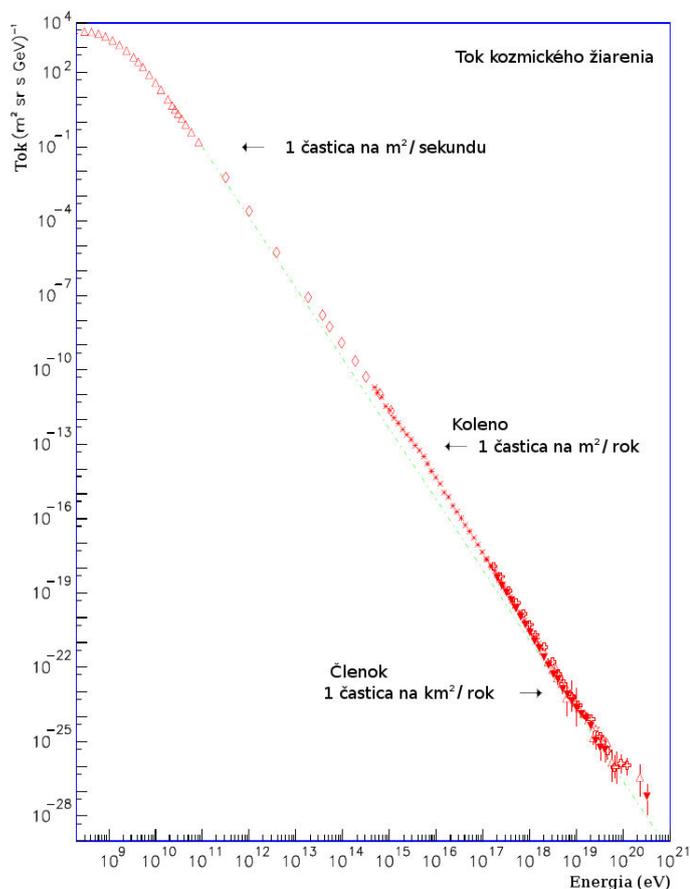
extenzívne časticové spršky, vďaka ktorým sa dajú študovať aj tie najenergetickejšie primárne častice, keďže ich priama detekcia je prakticky nemožná.

2.1 Energetické spektrum

Primárne častice kozmického žiarenia sa vo vesmíre vyskytujú v obrovskom rozmedzí energií. So vzrastajúcou energiou sa naše vedomosti o primárnych časticách znižujú kvôli nižšiemu počtu meraní a získaných dát. Častice nachádzajúce sa na vysokoenergetickom konci spektra sú predmetom intenzívneho výskumu. Pri energiách nižších ako 10^{10} eV, pozri obrázok 2–1 je možné sledovať antikoreláciu so slnečnou aktivitou, keďže kozmické žiarenie týchto energií je ovplyvnené medziplanetárnym magnetickým poľom[3].

Nad touto hranicou sú primárne častice natolko energické, že ich ovplyvňovanie lokálnymi podmienkami je zanedbateľné a závislosť toku sa redukuje na jednoduchý zákon $dN/dE \sim E^{-\gamma}$ čo hovorí, že ak má byť zachytené významné množstvo udalostí, tak sa musí merať na veľkej ploche s veľkým expozičným časom[7]. Do energie 10^{15} eV je spektrálny index konštantný a približne sa rovná $\gamma \approx 2.7$. Následne do energie 10^{18} eV sa spektrálny index mení na hodnotu $\gamma \approx 3.0$. Miesto v spektre, kde sa sklon spektra mení z hodnoty 2.7 na 3.0 sa označuje ako koleno, čo je vidieť na obrázku 2–1. Za hodnotou 10^{18} eV klesá tok pomalšie a to s faktorom $\gamma \approx 2.1$. Tento zlom sa nazýva aj členok, pozri 2–1. Je vidieť, že na veľkom rozsahu energií tok primárnych častíc sleduje tú istú závislosť, až na vyššie spomenuté zmeny v sklone spektra[3].

Koniec spektra (energie nad 10^{20} eV) je z fyzikálneho hľadiska najzaujímavejší a dnes najaktívnejšie skúmanou oblasťou. Pre zložitú čitateľnosť týchto energií je potrebné pre ich štúdium vynaložiť obrovské materiálne prostriedky. Medzi pozemné projekty, ktoré skúmali a skúmajú tieto vysoko energetické žiarenia patria japonský projekt AGASA[4], HiRes a jeho nasledovník Telescope Array[6] v Utahu a Pierre



Obr. 2–1 Spracovanie merania rôzneho spektra energií kozmického žiarenia. Bodkovaná čiara reprezentuje zákon mocniny E^{-3} . Taktiež sú uvedené integrálne toky[7].

Auger Observatory[5] v Argentíne. Štúdium vysokoenergetického konca spektra povedie k identifikácii zdrojov UHECR, ktoré sú zodpovedné za akceleráciu častíc k tak obrovským energiám[3]. Tým by bolo možné rozlúštiť hádanku mechanizmu akcelerácie najenergetickejších primárnych častíc kozmického žiarenia. Len pre ilustráciu, urýchľovač častíc LHC v CERNe dokáže urýchliť protóny na energiu 7×10^{12} eV[3].

2.2 GZK limit

V roku 1966 bola K. Greisenom a nezávisle na ňom aj G.T. Zatsepinom a V.A. Kuzminom navrhnutá idea, v súčasnosti známa ako GZK limit. Tento limit popisuje

správanie sa primárnych častíc s energiou väčšou ako 5×10^{19} eV. Pre tieto primárne častice prestáva byť medzihviezdny priestor transparentný a dochádza k ich brzdeniu zrážkami s fotónmi mikrovlnného žiarenia. Pod hranicu 5×10^{19} eV je interakcia s fotónmi elastická, a preto pre spektrum neviditeľná[8].

GZK limit predpokladá, že ak zaregistrujeme protón s energiou väčšou ako 5×10^{19} eV tak vzdialenosť zdroja, ktorý danú časticu takto urýchlil od miesta registrácie je menej ako 50 Mpc. Z toho vyplýva, že informácie o pôvode primárnych častíc je možné získať len vtedy, ak sú ich zdroje dostatočne blízko. A tu vzniká záhada. Na Zemi boli zaznamenané spršky, ktoré zodpovedajú primárnej častici s energiou väčšou ako 5×10^{20} eV, no v okolí 50 Mpc nie je známy zdroj, ktorý je schopný produkovať takéto častice[8].

2.3 Pôvod kozmického žiarenia

Podobnosť zloženia primárnych častíc a typického medzihviezdného prostredia naznačuje, že kozmické žiarenie je zvyčajná hmota, urýchlená k vysokým energiám. Za možné akceleračné mechanizmy už boli navrhnuté viaceré procesy. Kozmické žiarenie samotné by malo sledovaním ich trajektórie viesť späť k svojmu dominantnému zdroju. Lenže táto úvaha naráža na veľký problém. Počas svojej púte je primárna častica vystavená účinkom rôznych magnetických polí. Keďže medzihviezdne a medziplanetárne pole obsahuje nepravidelnosti, na ktorých sa trajektória kozmického žiarenia nižších a stredných energií mení, tak výsledná trajektória vyzerá ako difúzia v lokálnych magnetických poliach. Cesta konkrétnej častice nízkych a stredných energií má potom charakter náhodného pohybu a stráca sa informácia o polohe jeho zdroja. Trajektóriu častíc UHECR tieto nepravidelnosti, ani medzigalaktické a medzihviezdne magnetické pole v rámci GZK limitu, pozri obrázok 2.2, ovplyvňujú minimálne, a tak je možné sledovaním smeru ich príchodu určiť ich zdroj (platí pre energie minimálne $E \sim 10^{20}$ eV)[9].

2.4 Spřšky kozmického žiarenia

Pri zrážke primárnej častice s molekulou vzduchu atmosféry Zeme vzniká zárodok spřšky sekundárnych častíc. Ak je energia primárnej častice dostatočná, tak sa spřška dostane až k zemskému povrchu (EAS). Pod pojmom sekundárne častice sa predstavuje séria interakcií, kde v prípade, ak mala primárna častica dostatok energie, tak každá nová zrážka vytvára nové častice, ktoré produkujú ďalšie a spřška sa lavínovito šíri. Vzhľadom k zachovaniu hybnosti je smer spřšky totožný s pôvodným smerom primárnej častice a jej tvar pripomína kužel. Tak ako lavína, aj spřška sa v určitom bode zastaví. Celý proces od vzniku spřšky po jej ukončenie trvá len niekoľko mikrosekúnd[10].

2.5 Pozorovateľné efekty

Okrem priameho pozorovania sekundárnych častíc, vytvárajú tieto častice aj sprievodné javy - Čerenkov jav a fluorescencia dusíka.

- **Čerenkov jav** Pokiaľ sa nabitá častica šíri médiom rýchlejšie ako je rýchlosť svetla v tom prostredí, tak dochádza k Čerenkovmu javu. Pohybujúci náboj vytvára osciláciu atómov, ktoré produkujú elektromagnetické vlnenie. V momente, keď rýchlosť presiahne rýchlosť šírenia signálu, vlny začnú spolu interferovať a vytvoria rázovú vlnu čím vzniká žiarenie, ktoré je označené ako Čerenkovo žiarenie. V spřške sú za produkciu tohto žiarenia zodpovedné mióny[11].
- **Fluorescencia dusíka** Pri prechode atmosférou spôsobuje spřška sekundárnych častíc pozdĺž dráhy letu excitáciu atómov dusíka . Pomalý návrat týchto atómov z excitovaného stavu do základného stavu vedie k vzniku žiarenia v ultrafialovej a viditeľnej oblasti elektromagnetického spektra[12].

3 Sledovanie vysokoenergetického kozmického žiarenia

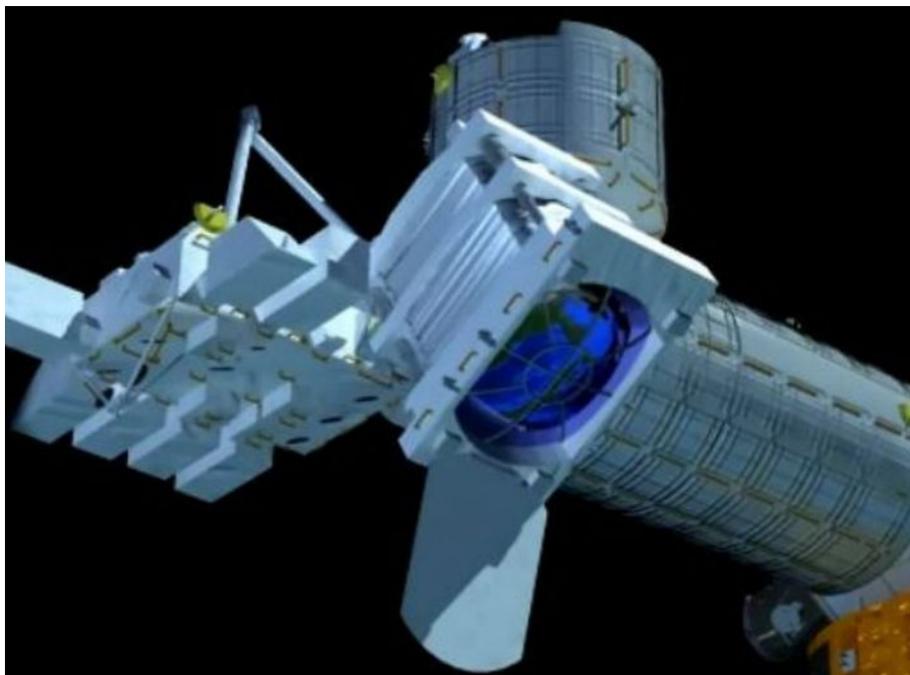
Kozmické žiarenie ultravysokých energií je doposiaľ sledované len pozemnými experimentmi ako je AUGER či Telescope Array. Na to, aby sme mohli zaznamenávať UHECR žiarenie na povrchu Zeme, musia existovať extrémne silné zdroje kozmického žiarenia urýchľujúce kozmické žiarenie na veľmi vysoké energie. Tieto zdroje sa musia nachádzať v obmedzenom rozsahu vzdialeností stanovenej GZK limitom[13].

Keďže distribúcia hmoty v rámci GZK limitu od Zeme je nehomogénna a anizotropná, dá sa očakávať, že smer príchodu UHERC bude zodpovedať tejto anizotropii. To sa zatiaľ nepotvrdilo dostatočne jasne. Na identifikáciu zdrojov UHECR sú potrebné merania, ktoré nám zobrazia energetické spektrum a smery príchodu UHECR na základe veľkého množstva registrovaných UHECR častíc. Vzhľadom na potrebnú štatistiku ide o náročnú požiadavku kvôli extrémne nízkemu toku kozmického žiarenia pri energiách $E > 5 \times 10^{19}$ eV, kde jedna častica dopadá na kilometer štvorcový povrchu atmosféry menej než raz za jedno storočie[13].

3.1 Experiment JEM-EUSO

Kozmické Observatórium Extrémneho Vesmíru na palube Japonského Experimentálneho Modulu (JEM-EUSO) na Medzinárodnej vesmírnej stanici ISS bude inovatívnou vesmírnou misiou. Najväčší rozdiel oproti doterajším metódam je v tom, že JEM-EUSO využíva zemskú atmosféru ako gigantický detektor[15].

JEM-EUSO detektor sleduje ultrafialové žiarenie produkované sprškami sekundárneho kozmického žiarenia. JEM-EUSO zachytí podstatne viac spršok ako podobne zamerané observatória na zemskom povrchu, keďže pri pohľade na Zem vidí omnoho väčšiu plochu než jej konkurenti na povrchu Zeme. Teleskop bude uchytený na ISS a to konkrétne na “exposed facility (EF)” na JEM module Kibo[15].



Obr. 3–1 Ilustračný pohľad na JEM-EUSO teleskop namontovaný na ISS.[14]

Hlavným cieľom experimentu je astronómia UHECR častíc. Medzi ďalšie ciele misie patrí detekcia vysokoenergetických gamma lúčov a neutrín. JEM-EUSO na pozorovanie UHECR častíc využíva fluorescenčné svetlo, ktoré je emitované počas vytvárania atmosférickej spršky sekundárneho žiarenia. Detegovaný svetelný signál spršky inicializovanej primárnou UHECR časticou v atmosfére sa využíva pre odhadnutie energie častice a smeru jej príchodu. Túto techniku využíva niekoľko pozemných observatórií, no technika pozorovania z vesmíru, z obežnej dráhy Zeme ešte nebola použitá[15].

3.2 Nástroje JEM-EUSO experimentu

Jednou z hlavných častí experimentu je teleskop, ktorý je citlivý na UV žiarenie. Optika teleskopu sa aktuálne skladá z dvoch zakrivených obojstranných frešnelových šošoviek s externým priemerom 2,6 metra a so stredne zakrivenou frešnelovou šošovkou. Kombinácia troch frešnelových šošoviek umožňuje plné využitie veľkého zorného

poľa s uhlom 60° s rozlíšením 0.075 stupňa, čo predstavuje asi 550 metrov zeme na jednom pixeli. Plocha detektora obsahuje 137 modulov detektorov fotónov, v skratke PDM, kde každé PDM sa skladá z 9 elementárnych častíc a každá elementárna častica obsahuje fotónový multiplikátor zložený zo 64 pixelov. Všetky tri frešnelové šošovky teleskopu JEM-EUSO je vidieť na obrázku 3–2[18].

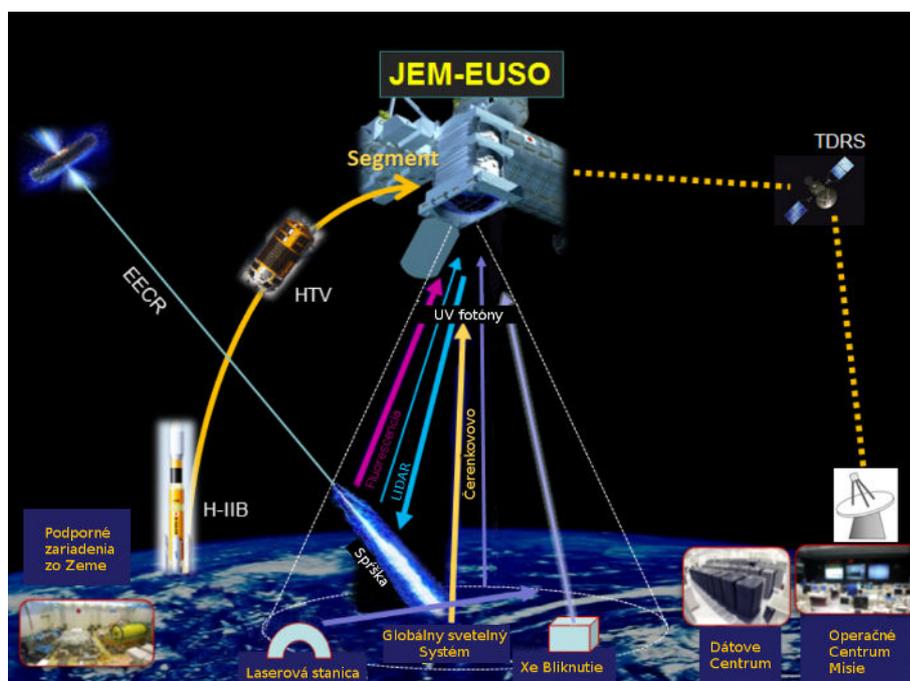


Obr. 3 – 2 Frešnelové šošovky..[16]

Snímacia rýchlosť teleskopu sa meria rádovo v mikrosekundách. Plocha detektora obsahuje 3×10^5 pixelov. Veľmi široké zorné pole kamery ($2 \times 30^\circ$) dovoľuje sledovať veľkú plochu na Zemi[18].

3.3 JEM-EUSO a rozpoznávanie obrazcov

JEM-EUSO diaľkovým prieskumom skúma fluorescenčné fotóny produkované EAS v atmosfére a sleduje pohyb stopy spŕšky (čela spŕšky) každých pár mikrosekúnd. Funguje ako extrémne vysokorýchlostná digitálna videokamera snímajúca UV žiarenie. Intenzita svetla pri nadol smerujúcej trajektórii atmosférickej spŕšky ukazuje energiu a smer príchodu UHECR udalosti kozmického žiarenia. Energia primárnej častice je úmerná výške maxima produkcie UV žiarenia v atmosfére[19]. Ilustráciu spôsobu fungovania JEM-EUSO experimentu je vidieť na obrázku 3–3.

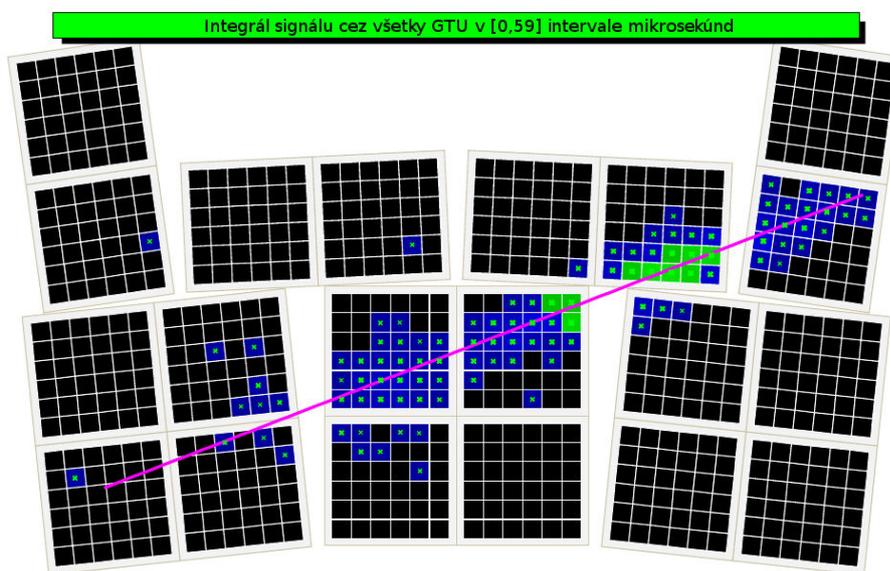


Obr. 3–3 Náčrt fungovania JEM-EUSO experimentu.[17]

Udalosti kozmického žiarenia sú zaznamenávané ako pixely v tvare dvojrozmerného poľa, ktoré budem ďalej nazývať snímky. Prístroj zachytáva až 400 000 snímok za sekundu. Pre zredukovanie počtu zachytených udalostí slúžia stupne spúšťačov. Ich úlohou je odstránenie udalostí, na ktorých určite nieje stopa spŕšky. Zvyšné snímky, ktoré prešli úrovňami spúšťačov, sú posielané na Zem, a na týchto fotkách sa budú hľadať potenciálne spŕšky[19].

4 Zachytenie spršky JEM-EUSO detektorom

Ako už bolo spomenuté, JEM-EUSO experiment sa snaží zaznamenávať UEHCR. Keď sprška s energiou 10^{20}eV natrafí na zemskú atmosféru, spôsobí fluorescenciu dusíka a Čerenkov jav, čo je opísané v kapitole 2.5. Pre zachytenie spršky sa použijú veľmi citlivé senzory na detekciu fotónov. Pole detektora obsahuje 3×10^5 pixelov. Na obrázku 4–1 je vidieť príklad zachytenia spršky detektorom.



Obr. 4–1 Ukážka distribúcie fotoelektrónov na ohniskovej vzdialenosti povrchu pre spršku s energiou 10^{20}eV a zenitový uhol $\theta = 60^\circ$. [33]

Pretože JEM-EUSO detektor bude zachytávať spršku vo veľkých rýchlostiach, je potrebné redukovať obrovské množstvo dát, ktoré vyprodukuje. Taktiež je potrebné rozlíšiť platné udalosti od šumu pozadia. Na rýchlu detekciu bola navrhnutá hierarchická spúšťacia schéma, ktorá pozostáva z prvého spúšťacieho stupňa, pozri sekciu 4.1, a druhého spúšťacieho stupňa, vid' sekciu 4.2. Schéma je postavená na rozdelení ohniskovej vzdialenosti povrchu do menších častí, nazývaných moduly svetelného detektora (PDM), nad ktorými sa vykonáva analýza. Rozdelenie spúšťacích stupňov je vidieť v tabuľke 4–1 [32].

Tabuľka 4–1 Jednotlivé stupne spúšťača a ich prislúchajúce časti [32].

Spúšťacie stupne	
Prvý spúšťací stupeň	Spúšťač detekcie fotónov
	Spúšťač súčtu fotónov
	Spúšťač spojitosti
Druhý spúšťací stupeň	

4.1 Prvá spúšťacia úroveň

Funkcia prvého spúšťacieho stupňa sa môže rozdeliť na tri menšie časti[32]:

- Sensory použité v detektore JEM-EUSO sú natoľko citlivé, že sú schopné zachytiť samotný fotón. Prvá časť spúšťača rozoznáva jednotlivé fotóny. Ak je fotón detekovaný, čiže hodnota nameraná na senzore je väčšia ako minimálna hranica, tak sa zavolá druhá časť prvého spúšťača.
- Druhá časť spúšťača pracuje na úrovni pixelov. Jej úlohou je počítať zachytené fotóny. Čo znamená, že ak je detekovaný fotón na jednom z detektorov, ktoré patria danému pixelu, tak pre tento pixel sa zvýši počet zachytených fotónov o 1. Na tejto vrstve sa dobre redukuje pozadie správnym nastavením hodnoty výšky minimálnej hranice nad kolísavou hodnotou pozadia.
- Tretia časť spúšťača sa vykonáva, ak v druhej časti je súčet fotónov na pixel väčší ako minimálna hranica a keď aktivita je trvajúca v po sebe idúcich jednotlivých zaznamenávacích častiach detektora. V tomto prípade sa robí sumácia zoskupených pixelov o veľkosti 2x2, alebo 3x3. Následná sumácia je porovnávaná s vopred definovanou hodnotou. Ak je hodnota sumácie menšia ako vopred definovaná hodnota na porovnanie, tak sa zresetujú počty fotónov pre dané pixely, ak je väčšia, tak sa spustí druhá spúšťacia úroveň.

4.2 Druhá spúšťacia úroveň

Spúšťač na druhej úrovni dáva finálne rozhodnutie pre štart procedúry na čítanie, respektíve posielanie dát na Zem na vykonanie analýzy. Spúšťač na druhej úrovni musí získať dáta z celej ohniskovej vzdialenosti povrchu, rozdelených na 18 častí skladajúcich sa z 8 PDM[34]. Tento spúšťač poskytne rozhodnutie o existencii možnej spršky zbieraním dát z prvého spúšťača. Algoritmus má názov spúšťač lineárnej dráhy (LTT), ktorý je určený na hľadanie svetelných bodov pohybujúcich sa rýchlosťou svetla 400 km dopredu.

Algoritmus si na začiatku definuje “skrinku” maximálneho počtu pixelov na GTU okolo pixela, nad ktorým sa spustil prvý spúšťač. Obsah vrámci “skrinky” je integrovaný. Pozícia skrinky v konkrétnom stálom GTU závisí na smere, v ktorom sa práve analyzuje. Pre ilustráciu, ak ide o vertikálnu spršku, tak “skrinka” bude pozostávať z tých istých pixelov v rôznych stálych GTU, zatiaľ čo pre horizontálnu spršku sa “skrinka” posúva v priemere o jeden riadok a stĺpec pre každé GTU [34].

Keď je už známa pozícia “skriniek” v stálych GTU, tak sa integrované hodnoty sumácie fotónov zo “skriniek” zosumujú. Tento výsledok sa porovnáva s aktuálnou minimálnou hranicou hodnoty požadovanej pre zachytenie spršky. Ak je hodnota väčšia ako hranica, tak sa dáta posielajú na Zem pre obsirnejšiu analýzu [34].

4.3 Udalosť určená na spracovanie

Po splnení podmienky aj na druhej spúšťacej úrovni, pozri sekciu 4.2, nastáva čítacia fáza. Tu sa v určitom okamihu pošle na Zem zachytená udalosť. Každá udalosť je zložená z 31 GTU, kde prvých 15 GTU sú dáta zachytené pred GTU, ktoré prešlo druhou spúšťacou úrovňou, čo je zároveň 16. GTU a posledných 15 GTU sú dáta zachytené po 16. GTU. Jednotlivé GTU predstavujú maticu o rozmeroch 36x36. Každý pixel tejto matice zahrňuje integrovaný počet fotoelektrónov, ktoré boli za-

chytené na JEM-EUSO detektore. Každá udalosť poslaná na zem prejde algoritmom pre rozpoznávanie vzorov, ktorý sa pokúsi nájsť spĺšku a určiť jej smer príchodu.

5 Rozpoznávanie vzorov

Jednoduchosť akou dokážeme rozpoznať písane slovo, odlíšiť vôňu cibule od vône ruže, rozlíšiť ženu od muža na základe kriviek zobrazuje neuveriteľné zložité procesy, ktoré sú základom rozpoznávania vzorov [20]. Naša schopnosť rozpoznávania sa rozvíjala desiatky tisíc rokov. V skratke, rozpoznávanie vzorov môžeme opísať ako vykonanie rozhodnutia na základe vstupných dátach.

Je prirodzené, že v dobe technológií sa ľudia snažia zostavovať stroje, ktoré sú schopné rozpoznávať vzory. Či už ide o identifikáciu otláčkov prstov, rozpoznávanie tvári, rozpoznávanie vzorov pomocou strojov môže byť pre ľudstvo veľmi užitočné. Implementáciou rozpoznávacích vzorov získavame hlbšie porozumenie princípov, na základe ktorých sa vytvárajú špeciálne algoritmy[21]. Rozpoznávanie vzorov je vedecká disciplína, ktorej cieľom je zaradenie objektov do množiny tried alebo kategórií.

5.1 Klasifikácia vzorov

Podľa Dudu [23] klasifikácia sa da rozdeliť na niekoľko podproblémov:

- **Extrakcia príznakov** Úlohou je oddelenie príznakov, na základe ktorých je možná klasifikácia objektu.
- **Šum** Všetky netriviálne rozhodnutia a problémy spojené s rozpoznávaním vzoru produkujú šum samé o sebe. Šum je vlastnosť snímaného vzoru, kde senzor nesníma len základný vzor, ale zachytáva aj náhodné dáta, šum, buď z okolitého sveta, alebo zo seba samého. Často sa využívajú predprocesy, ktorých úlohou je odstránenie šumu.
- **Štatistická odchylka** Príliš komplexný model môže spôsobiť výborné klasifikačné výsledky nad tréningovými dátami, no je veľmi málo pravdepodobné dosiahnuť podobné výsledky nad novými vzormi. Táto situácia sa

nazýva štatistická odchylka.

- **Selekcia vzoru** Určenie správnej triedy, do ktorej patrí vzor na základe jeho vlastností. Jednou z najťažších je rozhodnutie, kedy skúmaný vzor na základe vlastností máme zaradiť do inej triedy ako iné vzory s podobnými vlastnosťami.
- **Predchádzajúce znalosti** Aplikovanie už nadobudnutých znalostí môže pomôcť pri dizajne klasifikátora, kde sa použijú sľubné vlastnosti. Tu si vieme predstaviť ako špecifické atribúty formy, voľne povedané, vieme, že tvár ma dve oči jeden nos a podobne.
- **Chýbajúce príznaky** Pri klasifikácii vzoru sa môže stať, že nie je dostupný jeden z príznakov. Často sa to rieši tak, že chýbajúci príznak je buď nulový, alebo sa mu priradí priemerná hodnota získaná pomocou tréningových dát.
- **Mereológia** Zaoberá sa vzťahom medzi celkom a jeho časťami. Ako príklad si zoberme slovo teplomer. Rozpoznávač vzorov musí vedieť, kedy má rozpoznať len určitú časť vzoru, v tomto prípade teplo alebo celok, čiže teplomer.
- **Segmentácia** V praxi hľadané vzory nie sú ideálne izolované a pripravené na rozpoznávanie, no často sú zle otočené, prekrývajú sa. Systém musí rozpoznať, kde vzor začína a kde končí, inak povedané, vzory musia byť segmentované.
- **Kontext** Pre zlepšenie rozpoznávania vzoru sa môžu použiť kontextové údaje, dáta získané zo vstupu, ktoré sú odlišné od dát cieľového vzoru. Ak nám príde séria tých istých vzorov za sebou, je veľká pravdepodobnosť, že nasledujúci objekt, ktorý sa bude skúmať, patrí do tej istej triedy ako predchádzajúce skúmané vzory.

- **Nemennosť** Pri získavaní optimálnych výsledkov pri klasifikácii vzorov sa stretávame s problémom nemennosti. Pozícia snímaných objektov v rámci zaznamenávaného miesta nemusí byť vždy rovnaká. Na základe toho sa snažíme vytvoriť rozpoznávač, ktorý je nemenný voči transformáciám.
- **Združovanie údajov** Vylepšenie výsledkov rozpoznávania môžeme dosiahnuť použitím viacerých klasifikačných komponentov. Ak každý komponent súhlasí, tak vzor vieme zaradiť do triedy. Ak niektorý z komponentov nesúhlasí, spôsob, akým sa združujú údaje a následné rozhodnutie závisia od typu skúmaného objektu.
- **Risk a cena** Všeobecne sa po klasifikácii vzoru vykoná akcia zatriedenia skúmaného objektu na kopu. Avšak stále je tu risk, že pri výskyte chyby sa objekt zle zatriedi. Tu je cieľom minimalizovať percento nových vzorov, nazývaných ako zlá kategória.
- **Komplexita výpočtov** Je to meranie náročnosti výpočtu a škálovateľnosti algoritmu na základe rozmerov príznakov alebo počtu vzorov a kategórií.

5.2 Adaptácia a učenie sa

Na základe Polikára [22] je slovo učenie sa pri dizajne rozpoznávača získavanie akýchkoľvek informácií od tréningových dát. Pretože všetky zaujímavé alebo praktické rozpoznávania vzorov sú natoľko náročné, že nevieme jednoducho uhádnuť klasifikačné rozhodnutia, je potrebné stráviť veľké množstvo času učením. Vytváraním klasifikátorov sa vytvárajú predpoklady všeobecnej formy modelu alebo jeho tvaru a následne použitie tréningových vzorov na získanie neznámych parametrov modelu. Učenie reprezentuje algoritmus, ktorý sa snaží zminimalizovať chybovosť na skupine tréningových dát. Podľa Dudu [23] vieme učenie rozdeliť do niekoľkých častí:

- **Učenie pod dohľadom** Učiteľ poskytuje označenie kategórie alebo hodnotu

každého vzoru v tréningovej vzorke a vyhľadajú sa redukcie hodnôt pre vzor. Je tu snaha dosiahnuť, aby učiaci sa algoritmus bol dostatočne výkonný na zistenie parametrov pre riešenie hľadaného problému.

- **Učenie bez učiteľa** Učenie bez dohľadu, ináč aj zhlukovanie je metóda, kde sa vytvárajú zhluky alebo prirodzené skupiny na základe vstupných vzorov. Prirodzená skupina je vždy definovaná explicitne alebo implicitne v zhlukovacom systéme samom o sebe a dáva čiastočnú množinu vzorov, kde odlišné zhlukovacie algoritmy produkujú rôzne zhluky výsledkov.
- **Učenie so spätnou väzbou** Najtypickejšou cestou natrénovania klasifikátora je dať mu vstupné údaje, získať orientačné označenia kategórií a použiť už známe označenie cieľovej kategórie pre zvýšenie efektívnosti klasifikátora. Ak výsledok rozpoznávania znaku je napríklad písmeno R a požadované je B, tak v donucovacom učení neexistuje signál s očakávaným vstupom, ale je tam jednoduchý výstup či rozpoznanie je správne, alebo nie. V prídavnom učení vzorov je výsledok často v binárnej podobe buď je to správne, alebo nie.

5.3 Vyhodnotenie výsledkov

Len zriedkavo sú problémy spomenuté v sekcii 5.2 riešené osobitne. Čiže napríklad v snahe znížiť zložitosť klasifikátora môžeme ovplyvniť jeho schopnosť vyrovnávania sa s nemennosťou[24].

5.4 Zhluková analýza

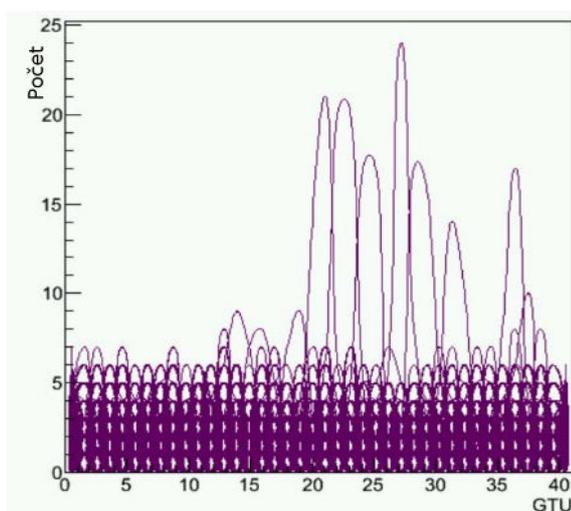
Zhluková analýza je zaužívaná technológia, ktorá sa okrem iného využíva aj pri rozpoznávaní vzorov[25]. Vzor je predloha vlastností a ich hodnôt na základe ktorých definujeme objekt. Zmenou hodnôt vlastností sa môže zmeniť aj definovaný objekt.

Ak napríklad medzi vlastností vzoru pre rozpoznanie motorového vozidla patrí dĺžka a šírka, tak zmenou hodôt sa môže zmeniť definovaný objekt z osobného auta na nákladne auto.

Zhluková analýza triedi analyzované jednotky do zhlukov tak, aby prvky toho istého zhukov si boli podobnejšie ako prvky z iných zhlukov. Zhlukovanie nie je špecifickým algoritmom, ale opisuje všeobecnú úlohu, ktorú treba riešiť. Výsledok môže byť dosiahnutý viacerými algoritmi, ktoré sú odlišné tým, čo pre nich znamená zhuk a ako ho efektívne nájsť[25].

5.5 PWISE metóda na vyhľadávanie spršok

Pre úspešnú rekonštrukciu udalosti musí byť signál výrazne oddelený od pozadia. Jedným z algoritmov, ktorý sa o to pokúša je algoritmus “Technika pre vyhľadávanie okien a maxima”, skrátene PWISE [26], vyvíjaný členmi kolaborácie JEM-EUSO. Základnou myšlienkou PWISE metódy je, preskúmanie celého časového priebehu signálu (nie šumu) na každom pixeli. Môžeme to prirovnať k osciloskopickému zobrazeniu spriehu signálu na každom pixeli, vid’ obrázok 5–1, ktorý zaznamenal nárast



Obr. 5–1 Znáznorenie šumu pozadia s maximálnou hodnotou 5 zmiešaného s hodnotami signálu, získaných metódou PWISE.[27]

a pokles signálu a horná hranica by sa vytiahla z hodnôt registrovaného pixelu, čo je len signál, šum je tu odstránený.

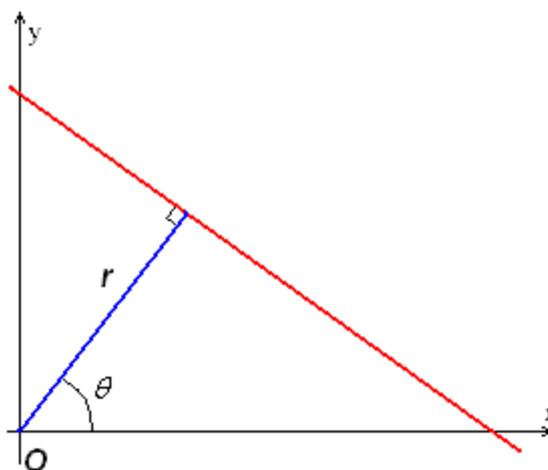
Podľa [26] je algoritmus PWISE metódy možné rozdeliť do troch krokov.

- **Krok prvý** Algoritmus prechádza každým pixelom a vyberá spomedzi nich len tie, ktorých hodnota, tj. počet registrovaných fotónov, je väčšia ako je hodnota prahovej hodnoty. Táto bola na základe porovnávania signálu simulovaných spíšok a šumu pozadia stanovená na hodnotu 8 a tvorí prvý parameter metódy PWISE. Pre pixely spĺňajúce túto podmienku si zapamatáme hodnotu pixelu a jeho GTU.
- **Krok druhý** Pre každý pixel, ktorý splnil prvú podmienku, sa určí séria hodnôt odstupu signálu od šumu, v skratke SNR. Séria SNR hodnôt sa určuje pre množinu rôzne dlhých časových okolí delta Tau pre vývoj signálu na vybranom pixeli v okolí jeho GTU. SNR, určíme ako $SNR = \frac{1}{\Delta\tau * RMS} * \sum_{\Delta\tau} Px(t)$,
 kde $RMS = \sqrt{\frac{1}{N} * \sum_{\Delta\tau} Px(t)}$, $\Delta\tau = N * GTU$.
 Množina hodnôt $\Delta\tau$ sa určuje okolo vybraného pixelu a jeho GTU. To znamená, že pre vybraný pixel a čas Δt registrácie nadprahovej hodnoty na ňom sú to tieto časové okná: $t \pm (0 * GTU)$, $\Delta\tau$ je jedno GTU,
 $t \pm (1 * GTU)$, $\Delta\tau$ je tri GTU,
 $t \pm (2 * GTU)$, $\Delta\tau$ je 5 GTU.
 a tak ďalej. Inak povedané, pre každý pixel sa určuje SNR vo viacerých časovo symetrických oknách okolo GTU s nadprahovou hodnotou z prvého kroku.
- **Krok tretí** Zo série odstupov signálu od šumu sa vyberie maximum. Danému SNR_{max} je pripisované $\Delta\tau SNR_{max}$. Ak je SNR_{max} väčšie ako hodnota 5, tak pixel prešiel PWISE filtrom, čo znamená, že bola splnená aj druhá podmienka.

PWISE filtráciou prejde pixel, ktorý súčasne spĺňa prvú aj druhú podmienku. Pre takýto pixel sú výstupom PWISE metódy hodnoty signálu na ňom v časovom okolí $\pm\Delta\tau SNR_{max}$ okolo času kde splnil prvú podmienku PWISE metódy.

5.6 Houghova transformácia pre všeobecné použitie

Houghova metóda vyhľadáva vzory pozdĺž priamej línie (vzoru). Cieľom Houghovej transformácie je získať inštancie objektov, ktoré su nedokonalé a patria špecifickej triede pomocou hlasovacej procedúry. Hlasovanie je vykonávané v priestoroch parametrov, z ktorých sa získavajú kandidáti objektov ako lokálne maximá. V praxi to znamená, že pre každý pixel prehľadávanej matice (PDM) určuje kolmú vzdialenosť k počiatku súradnicovej sústavy (zvyčajne pixel (0,0)), pre sériu línií ním prechádzajúcich, pozri obrázok 5–2. Pre maticu NxN pixelov pre každý pixel pre M smerov určuje NxNxM kolmých vzdialeností r. Množina týchto vzdialeností r tvorí takzvaný Houghov priestor.



Obr. 5–2 Ukážka priamej línie (červená čiara) a jej kolmej vzdialenosti k počiatku súradnicovej sústavy (modrá čiara).[28]

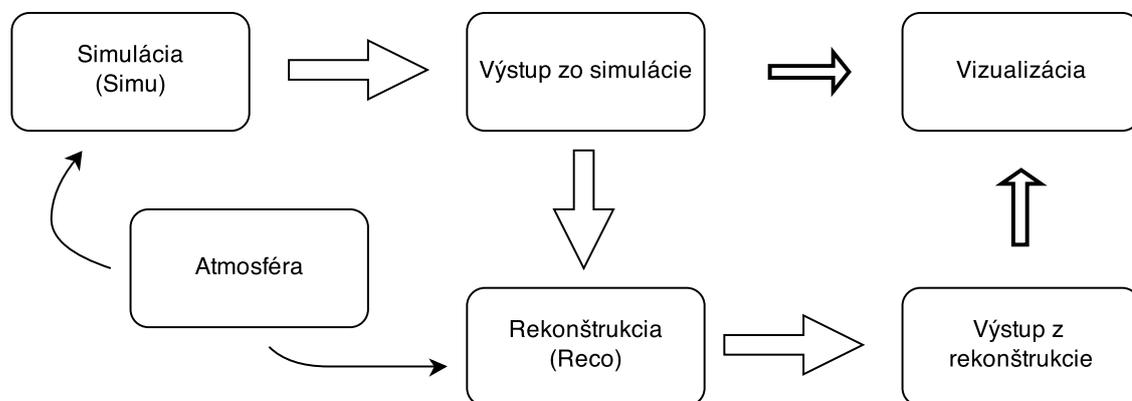
Pixely tvoria jednu líniu, respektíve nachádzajúce sa na nej, či v jej blízkosti, tvoria hľadaný priamy vzor. Počet a dĺžka nájdených vzorov závisí na definícii vzoru, res-

pektíve inak povedané, na parametroch, podľa ktorých sa v Houghovom priestore zoskupujú pixely na jednej línii (s podobnou či rovnakou kolmou vzdialenosťou r).

Jedna z najjednoduchších definícií vzoru, ktorá je natívne vlastná Houghovej metóde je nespojitý vzor. Tieto nespojité vzory hľadá Houghova metóda bez prídavných podmienok. Pre hľadanie spojitých podmienok musí byť Houghova metóda upravená. Združovanie bodov Houghovho priestoru pre nespojitý vzor je najjednoduchšie riešiť rozdelením možných hodnôt r na sériu rovnakých subintervalov. Ak sa r vyjadruje v jednotkách dĺžky strany pixelu, tak zaokrúhlením hodnoty r na celočíselné hodnoty sa zoskupia body Houghovho priestoru okolo vzdialeností $r \pm dr$ kde $dr = (\text{dĺžka hrany pixelu} / 2)$. To znamená v páse širokom jeden pixel. Počet nájdených vzorov závisí aj od počtu smerov M , a teda od delenia všetkých smerov podľa rovnakého uhla $\alpha = 360^\circ / M$.

6 Kolaboračný rámec projektu ESAF

Rámec pre simuláciu a analýzu je vyvíjaný kolaboráciou JEM-EUSO na to, aby poskytol vedcom koncový simulačný softvérový rámec. Pokrýva celý proces simulácie a analýzy dát od interakcie primárnej častice v atmosfére, po rekonštrukciu prípadu udalosti. ESAF je napísaný v jazyku C++ a FORTRAN použitím objektovo orientovaného prístupu. Rámec je založený na softvérovom balíku ROOT, viď sekciu 6.3 a taktiež využíva knižnice ROOTu. Na komunikáciu sa používa CINT, interpretér jazyka C++. Softvér ESAF je vysoko modulárny a skladá sa z niekoľkých podprogramov, z ktorých každý je zodpovedný za simuláciu špecifickej úlohy[29]. To znamená, že simulácia spršky, emisie svetla, propagácia fotónov v atmosfére a detektor, reakcie elektroniky a nakoniec rekonštrukcie udalostí, sú vyhodnotené nezávisle na sebe. Keď sa zavolá potrebný modul, zdedí všetky potrebné dáta na výpočet od predchodcu a po vykonaní výpočtu prepošle dáta nasledujúcemu modulu, kde sa susediace moduly správajú ako čierne skrinky. Modulárna štruktúra zaručuje jednoduché zmeny v komponentoch, pokiaľ sú zachované pôvodné rozhrania pre komunikáciu. To má za výhodu, že ESAF môže byť prispôsobený pre takmer ľubovoľnú úlohu[29]. Na obrázku 6–1 je vidieť rozdelenie softvérového rámca ESAF. Medzi hlavná časti ESAF rámca zaraďujeme časť simulácie, viď sekciu 6.1, a časť rekonštrukcie, pozri sekciu 6.2.



Obr. 6–1 Zobrazenie základného toku rámca ESAF.[33]

6.1 Simulačný balíček rámca ESAF

Simulácia rozdelená na 6 hlavných sekcií. Simulácia spršky začínajúca od toku vysokoenergetických kozmických častíc, produkcia fluorescenčných a červenkových fotónov v atmosfére, propagácia svetla od produkcie bodu smerom k obiehajúcemu teleskopu so simuláciou prebiehajúcich interakcií fotónov počas samotnej propagácie, simulácia optiky teleskopu a simulácia elektroniky a jednotlivých úrovni spúšťačov [33].

Fenu [30] zaraďuje medzi najdôležitejšie časti simulácie:

- **Spršky** Pri simulácii spršiek je dostupných niekoľko možností. ESAF poskytuje rozhranie podporujúce prístup už k vopred vygenerovaným sprškám simulovaným pomocou algoritmov ako sú CORSICA, CONEX alebo UNISIM. Najviac realistickú simuláciu spršky ponúka kód Monte Karlo CORSICA. Avšak dôsledkom je časovo veľmi náročný kód, preto nie je vhodný na štatistickú prácu v pomerne krátkom čase. Preto existujú zvyšné algoritmy, aby sa programátor mohol rozhodnúť či je potrebná maximálna realnosť, alebo väčší objem spršok za relatívne kratší čas.
- **Produkcia a propagácia svetla v atmosfére** Atmosférické podmienky sú nevyhnutné pre správnu simuláciu vzduchovej spršky, simuláciu fluorescenčného a Čerenkovho svetla a správne šírenie fotónov smerom k teleskopu. Taktiež rekonštrukčný modul musí byť oboznámený o aktuálnom stave atmosféry.
- **Detektor** Simulácia detektora je rozdelená do dvoch častí. Prvý modul je zodpovedný za prenos fotónov cez optiku nachádzajúcu sa na ohniskovom povrchu detektora, zatiaľ čo druhý modul simuluje správanie sa elektroniky, viď 4.

Výstupom simulačného balíka je ROOT súbor so štruktúrou, ktorá obsahuje tie isté informácie, ktoré so sebou nesú reálne dáta a ešte naviac skupinu "Monte Carlo

reálnych” dát. Úroveň detailov informácií je konfigurovateľná používateľom, aby bolo možné optimalizovať veľkosť výsledkov tak, aby to pokrylo ciele simulácie. Poskytnúť používateľovi ľahký prístup k dátam pomocou skupiny pohľadov na zdetekované udalosti. Pohľady nám umožňujú koherentné zobrazenie informácií aktuálneho výpočtu, histogramy a niekoľko 2D a 3D pohľadov vývoja udalostí v detektore a na ohniskovej ploche[33].

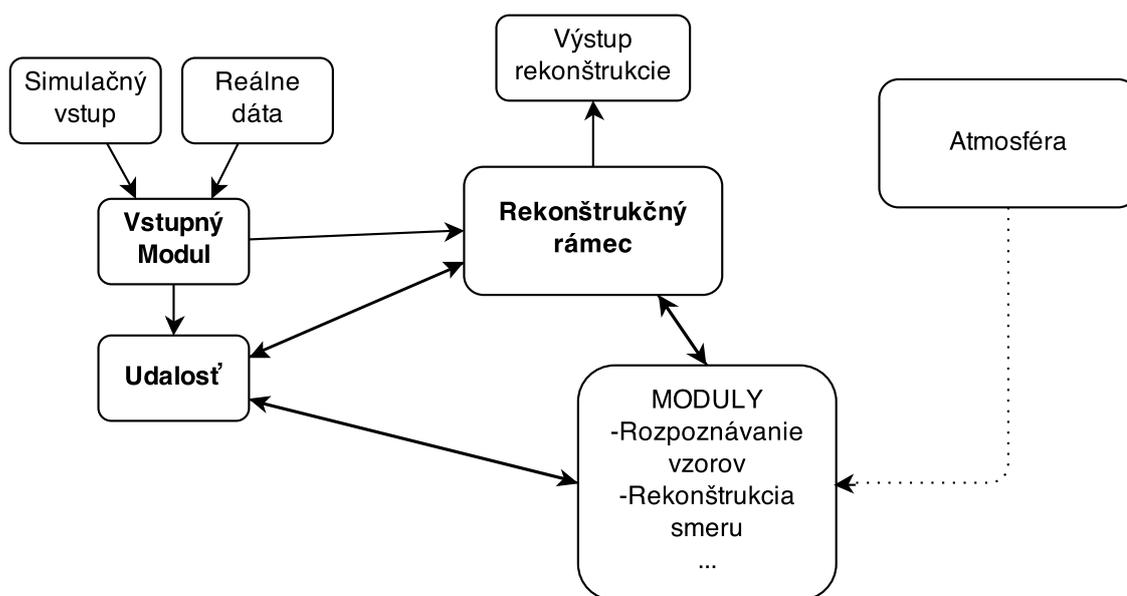
6.2 Rekonštrukčný balíček rámca ESAF

Fenu [30] hovorí, že rekonštrukčný model číta dáta späť z ROOT súbora, ktorý reprezentuje plne simulovanú spršku s pozadím a rekonštruuje udalosť tak, akoby išlo o reálne dáta z teleskopu. Niektoré z modulov sú zodpovedné za rozpoznávanie vzorov a identifikáciu EAS signálu nad nekoherentným svetlom z hviezd a pozadia zo svetla mesiaca, 3D rekonštrukciu smeru príchodu, energiu a rekonštrukciu maximálnej interakcie vo výške. Rekonštrukčný rámec je hlavnou štruktúrou, ktorá získava dáta (simulované alebo reálne) zo vstupného modulu a vytvorí reťazec modulov potrebných na zrekonštruovanie udalosti. Podľa Mernika [29] je samotná rekonštrukcia rozdelená do niekoľkých častí, kde pre každú z tých častí môže byť dostupný vlastný model.

- **Rozpoznávanie vzorov** Tento rámec sa zameriava na diskrimináciu reálnych udalostí od tých falošných generovaných náhodnými výkyvmi pozadia, svetlami miest alebo inými zdrojmi silného svetla pomocou odlišných modulov.
- **Rekonštrukcia smeru príchodu** Pre odhad prichádzajúceho smeru spršky primárnej častice existuje niekoľko modulov, ktorými sa to dá zistiť.
- **Energetická rekonštrukcia** Predtým ako môže byť energia vypočítaná je nevyhnutná znalosť výšky, kde sprška dosahovala najväčšiu energiu. Na získanie maximálnej výšky energie sú v ESAFe implementované dve tech-

niky, kde prvá na výpočet používa tvar svetelnej krivky a druhá Čerenkove echo odrážané od zeme.

Vďaka štruktúre rekonštrukčného balíka, pozri obrázok 6–2, sa dá jednoducho nahradiť, alebo odstrániť modul z rekonštrukcie. Týmto spôsobom vieme testovať a porovnávať rôzne algoritmy alebo ich kombinácie na tej istej udalosti. Tak ako v simulačnej časti aj moduly v rekonštrukčnej časti môžu byť konfigurované používateľom. Kontajner pre udalosť uchováva všetky dôležité informácie rekonštruovanej udalosti. Každý modul má prístup ku kontajneru udalosti, kde môžu čítať údaje zapísané v kontajneri a zapísať tam vlastné dáta[33].



Obr. 6–2 Zobrazenie štruktúry rekonštrukčného modelu rámca ESAF.[33]

6.2.1 Moduly v rekonštrukčnom balíku rámca ESAF

Na základe štruktúry rekonštrukčného balíka, viď sekciu 6–2, sa pridaná funkcionálnosť rieši pomocou modulov. To znamená, že či ide o rozpoznávací alebo rekonštrukčný algoritmus, sú pridávané vo forme modulov. Každý nový modul musí dedič od triedy s názvom RecoModule. Táto trieda modulov ponúka základné rozhranie

a funkcionálnu pre načítanie modulu rámcom ESAF a jeho zapojenie do analýzy sprášky. Programátor má možnosť prepísať zdedené funkcie a pridať vlastnú funkcionálnu. Rámec vždy pre každý modul zavolá ako prvé pred procesnú metódu, kde sa zvyčajne zvyknú načítať dáta a hodnoty premenných z konfiguračných súborov. Po tomto kroku s názvom inicializačný stav sa zavolá hlavná metóda modulu. Jej názov v prekladne znamená spracuj a ako parameter dostane sprášku, nad ktorou sa má samotná analýza vykonať. Následne rámec ESAF po spracovaní sprášky zavolá metódu na uloženie výsledkov vlastnej analýzy do výslednej spracovanej sprášky rekonštrukčného balíka a nakoniec sa zavolá metóda na očistenie pamäte od použitých lokálnych zdrojov.

6.2.2 Reprezentácia sprášky v softvérovom rámci ESAF

Po ukončení simulačného modulu sa dáta vygenerovaných sprášok uložia do súboru v rámci ROOT štruktúry, vid' sekciu 6.3, ktoré následne rekonštrukčný model prečíta a uloží si ich do pamäte pod svojou vlastnou štruktúrou. Dáta reprezentujúce sprášku obsahujú dve dôležité mapy jazyka C++, ktoré sa používajú na rekonštrukciu a rozpoznávanie vzorov. Prvou mapou štruktúry je

$$map < Int_t, vector < RecoPixelData* >> fPixelsGTU,$$

ktorá mapuje jednotlivé GTU k vektoru pixelov, ktoré prešli prvou a druhou úrovňou spúšťačov, pozri sekciu 4. Štruktúra reprezentujúca pixel obsahuje identifikačné číslo, počet zachytených fotoelektrónov, ich typ a mnoho iných informácií. Zo štruktúry reprezentujúcej sprášku sa vedia vytiahnuť aj údaje o geometrii detektora, kde sa napríklad na základe identifikačného čísla pixelu vie určiť jeho pozícia v milimetroch na ploche detektora. Ďalšou dôležitou mapou

$$map < Int_t, map < Int_t, vector < RecoPixelData* >>> fCellGTU,$$

ktorá mapuje PDM, ktoré prešlo prvou a druhou úrovňou spúšťačov na mapu pixelov a GTU, čo je opísané vyššie.

6.3 Softvérový balík ROOT

Na základe ROOT tímu[31] systém ROOT je sada objektovo orientovaných rámcov so všetkými funkciami potrebnými na spracovanie a analýzovanie obrovského množstva dát veľmi efektívnym spôsobom. Vďaka definovaniu dát ako skupiny objektov, sú použité špeciálne metódy pre priamy prístup k rozdielným atribútom vyznačených objektov bez priamej manipulácie priamo nad dátami. Rámec zahŕňa metódy pre vykresľovanie histogramov bez obmedzenia rozmeru, zrovnávanie kriviek, vyhodnocovacie funkcie, grafy a vizualizačné triedy pre jednoduché konfigurovanie analyzačných parametrov buď dávkovo, alebo interaktívne a v neposlednom rade rámec pre všeobecné paralelné výpočty.

Vďaka vstavanému CINT, interpretéru C++ jazyka, príkazového jazyka sú skriptovanie alebo vytváranie makier a programovanie vykonávané pomocou jazyka C++. Interpretér umožňuje rýchle zostavovanie makier, pretože odstraňuje časovo náročný linkovací cyklus. Systém bol navrhnutý spôsobom, že je možné sa dotazovať jeho databázy paralelne na výpočtových pracovných staniciach alebo mnohojadrových strojoch. ROOT je otvoreným systémom, ktorý môže byť dynamicky rozšírený pripojením externých knižníc podľa ROOT tímu[31].

Rámec ROOT vznikol ako požiadavka už na zastaralé knižnice PAW, PIAF, GEANT napísaných v jazyku FORTRAN. Aj keď tieto knižnice sú ešte stále populárne, nie je možné ich efektívne škálovať na požiadavky, ktoré vyžaduje napríklad Veľký hadrónový urýchľovač. Ponúka základný rámec so sadou funkcií a nástrojov pre všetky oblasti fyziky výpočtov vysokých energií[31].

7 Metódy skladania udalosti v 2D

Finálne rozpoznávanie spršky nad množinou udalostí, ktoré boli poslané na Zem z detektora JEM-EUSO má na starosti algoritmus pre rozpoznávanie vzorov. Tento algoritmus musí správne rozanalyzovať udalosť na základe hodôt pixelov v jednotlivých GTU. Taktiež musí brať na vedomie poradie GTU. Pre výpočet smeru príchodu spršky si je potrebné upraviť udalosť pozostávajúcu z 31 GTU tak, aby algoritmus pre rozpoznávanie vzorov bol aplikovaný nad finálnou, vyskladanou maticou o rozmeroch 36x36 z 31 GTU. Pre návrh a realizáciu algoritmu sa použije metóda klastrovania, viď sekciu 5.4, a jej otestovanie správnosti je možné vidieť v sekcii 8.2. Vo všeobecnosti sa môže algoritmus rozdeliť na dve základne časti:

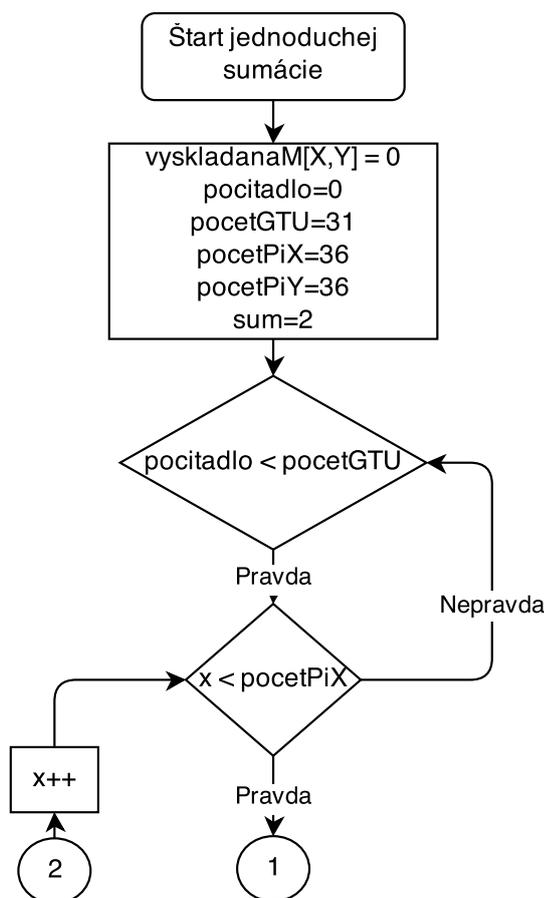
- algoritmus skladania udalosti,
- algoritmus pre rozpoznávanie vzorov.

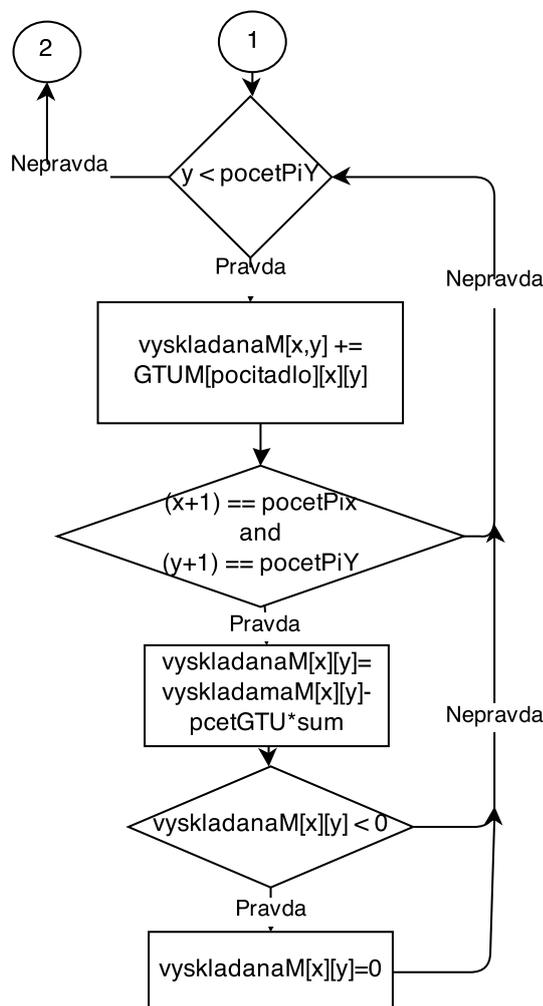
7.1 Návrh algoritmu pre vyskladanie matice

Každá udalosť, ktorá má byť analyzovaná sa skladá z 31 GTU. Je potrebné vytvoriť algoritmus, ktorý nám z týchto 31 GTU vytvorí vyskladanú maticu s rozmermi GTU. Ideálne algoritmus úplne odstráni nežiadúci šum z pozadia a zanechá len samotnú spršku. Každé z GTU v udalosti obsahuje 36x36 pixelov. Tieto pixely predstavujú počet zachytených fotoelektrónov detektorom JEM-EUSO. Pri vyskladávaní si určíme priemernú hodnotu šumu. Všetko nad touto hodnotou pre nás bude predstavovať potencionálnu spršku. Na začiatok aplikujeme jednoduchú sumáciu všetkých GTU nachádzajúcich sa v udalosti, pozri sekciu 7.1.1.

7.1.1 Vyskladanie matice jednoduchou sumáciou

Hlavnou myšlienkou algoritmu je jednoduchá sumácia hodnôt pixelu z každého GTU do finálnej vyskladanej matice. Ak je potrebné vypočítať hodnotu pixelu na pozícii 0,0 vo finálnej vyskladanej matici, tak sa sčítajú všetky hodnoty pixelov zo všetkých GTU na danej pozícii. Ak náhodou na vybranej pozícii nebude žiadny pixel s hodnotou väčšiou ako je vopred určená hranica šumu je možné, že sumáciou hodôt nie nulových no zároveň menších ako je hranica šumu bude súčet natoľko veľký, že sa bude javiť ako potencionálna spŕška. Preto vo finálnej matici od hodnoty každého pixelu odpočítame hodnotu $N=(\text{počet GTU v udalosti}) * (\text{hranica šumu})$. Ak výsledok po odčítaní bude záporný, tak nastavíme hodnotu pixelu na 0. Algoritmus vyskladávania je zobrazený na obrázku 7-1.

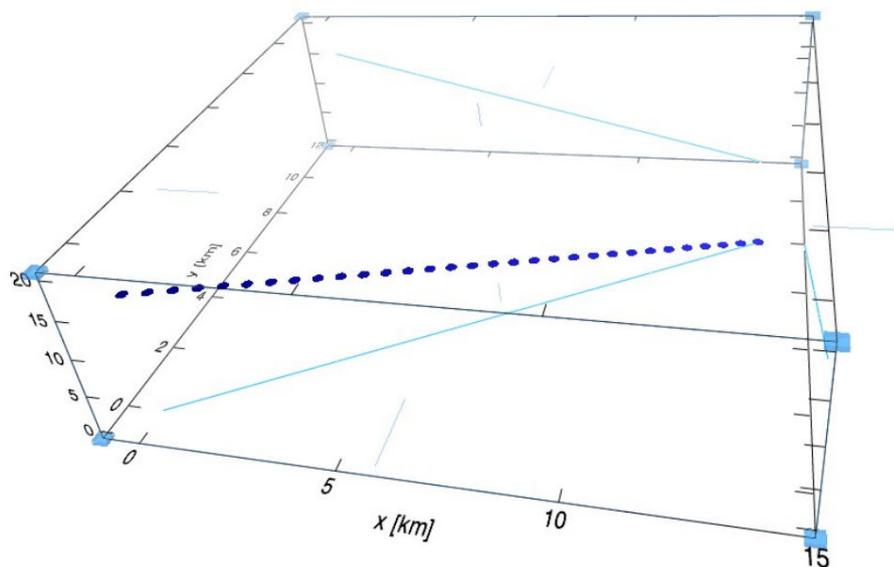




Obr. 7 – 1 Diagram postupov pre jednoduchý sumačný algoritmus.

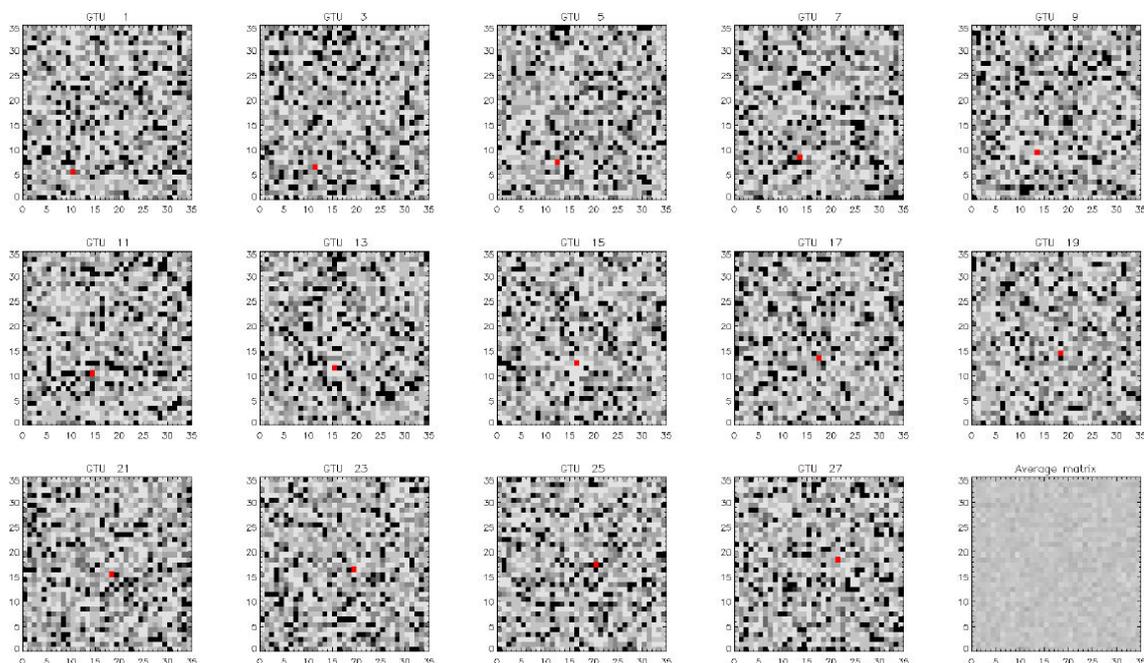
Na otestovanie vyššie spomenutého algoritmu si do udalosti, ktorá bude pokrytá šumom s Poissonskou distribúciou vložíme falošnú spršku. Falošná sprška sa vloží pod zenitovým uhlom 45 stupňov a pod axiálnym uhlom 30 stupňov. Signál generovaný sprškou sa uloží ako hodnota pixelu, ktorá je rovná súčtu hodnoty šumu a hodnoty signálu, čo bude 5. Taktiež prvá viditeľná interakcia bude na 15. kilometri, kde hodnoty začiatku spršky je vidieť na pozícii 0,0. Názorná ukážka spršky je na obrázku 7–2.

Po vložení falošnej spršky do testovacej udalosti sú pixely ožiarené sprškou označené



Obr. 7–2 Zobrazenie falošnej spřšky, ktorú vložíme do udalosti pre otestovanie algoritmu.

červenou farbou, na obrázku 7–3 pre lepšiu ilustráciu. Je potrebné podotknúť, že šum tu nebol preúrovňovaný a taktiež spřška nie je simulovaná, ale na základe jej príchodových uhlov vložená na správne miesta v udalosti pre ilustráciu princípu. Vpravo dole na obrázku 7–3 sa nachádza vyskladaná finálna matica podľa jednoduchého sumačného algoritmu. Ako je vidieť falošná vložená spřška nie je rozoznatelná voľným okom. Táto jednoduchá sumačná metóda úplne zahltila vloženú falošnú spřšku tak, že ju nie je možné rozoznať od šumu. Preto je potrebné modifikovať jednoduchý sumačný algoritmus, vid' 7.1.2.

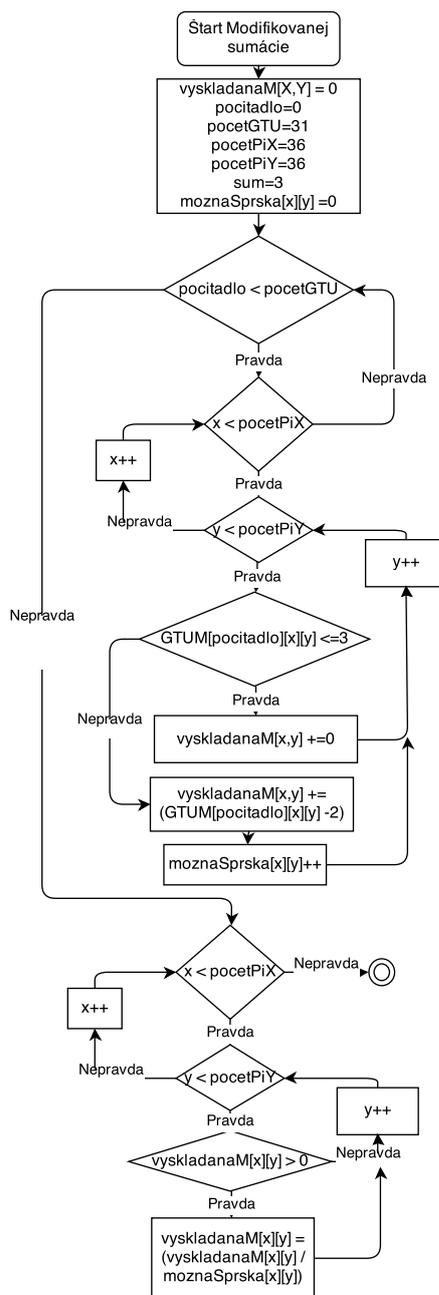


Obr. 7–3 Ilustračné zobrazenie vlozenej falošnej spršky do udalosti a zosumarizovanej matice podľa jednoduchovej sumácie.

7.1.2 Vyskladanie matice upravenou sumačnou metódou

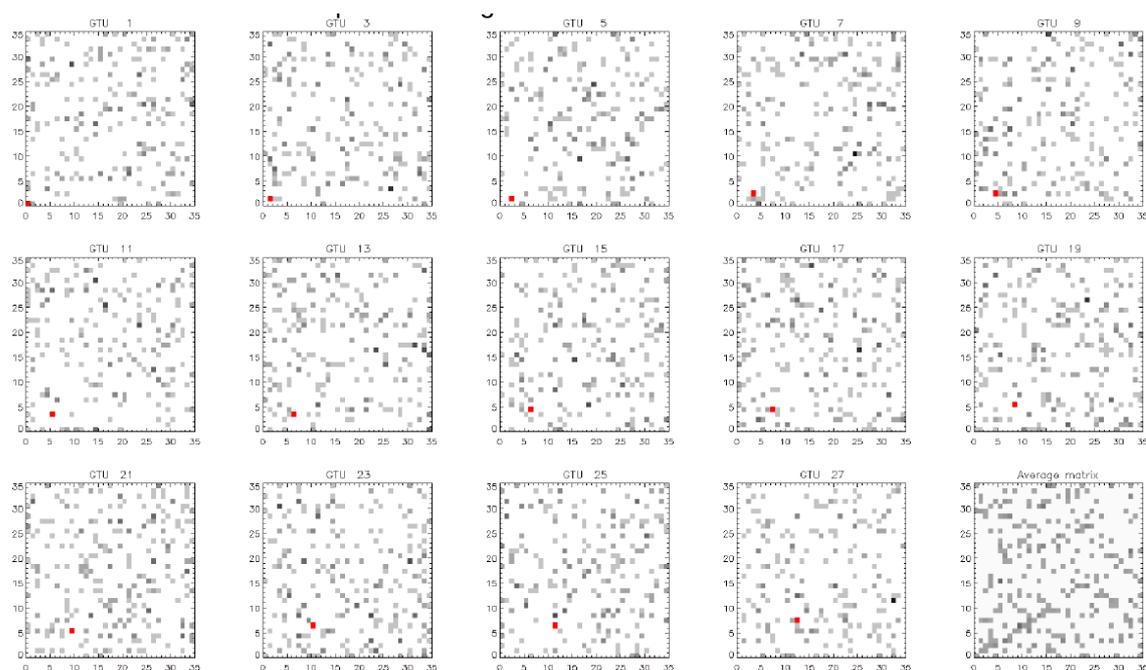
Jednoduchá sumačná metóda 7.1.1 nie je dostatočná na predprípravu udalosti pre rozpoznávanie spršky. Táto metóda by teoreticky fungovala pri menšom počte GTU. Pri aktuálnom počte sa pixely osvietené sprškou spriemerujú s hodnotami šumu z iných GTU, na ktorých sa sprška nachádza na odlišnej pozícii (záleží od nábehového uhla). Ako prvý krok je potrebné odstrániť nežiadúci šum z jednotlivých GTU, a potom na takto upravené GTU modifikovať sumačný algoritmus. Pre očistenie šumu pozadia sa jednoducho povie, že ak je hodnota pixelu GTU menšia alebo rovná šumu, tak pre daný pixel sa nastaví hodnota na 0. Ak je hodnota väčšia ako hranica šumu, tak sa od daného pixelu odpočíta hodnota šumu a nezabudne sa, že pre pozíciu daného pixelu sa redukoval signál od šumu pozadia. Takto sa upraví každé GTU z udalosti. Potom nasleduje trošku upravená sumácia. Hodnoty pixelov vo vyskladanej matici sa vypočítajú nasledovne. Vyberie sa pozícia pixelu, pre ktorý sa

spravi modifikovaná sumácia. Tam sa zo všetkých 31 GTU sčítajú hodnoty pixelov na danej pozícii. Výsledná hodnota sa potom predelí počtom, koľkokrát na tejto pozícii v rámci 31 GTU bola hodnota väčšia ako šum pozadia. Algoritmus je vidieť na obrázku 7–4.



Obr. 7–4 Diagram postupov pre modifikovaný sumačný algoritmus.

Na test algoritmu modifikovanej sumačnej metódy sa použije tá istá falošná sprška ako sa použila aj pri jednoduchom sumačnom algoritme, viď obrázok 7–2. Signál generovaný sprškou sa uloží ako hodnota pixelu, ktorá je rovná súčtu hodnoty šumu a hodnoty signálu, čo bude 5. Prvá interakcia bude na 15. kilometri. Vložená sprška je v jednotlivých GTU udalostiach pre lepšiu ilustráciu označená červenou farbou, viď obrázok 7–5. Finálna vyskladaná matica sa nachádza dole v pravom rohu na obrázku 7–5. Táto modifikovaná metóda prináša lepšie výsledky oproti jednoduchej sumácii. Vloženú spršku je možné pozorovať voľným okom. Je potrebné podotknúť, že táto sprška nebola simulovaná a pre test skladania bola vložená do udalosti, aby odpovedala prislúchajúcim uhľom dopadu.



Obr. 7–5 Ilustračné zobrazenie vlozenej falošnej spršky do udalosti a zosumarizovanej matice podľa modifikovanej sumácie.

Finálne matice, ktoré vznikli vyskladávaním GTU z udalostí pomocou modifikovaného sumačného algoritmu budú použité spolu s algoritmom pre rozpoznávanie potencionálne zachytenej spršky JEM-EUSO detektorom.

7.2 Návrh algoritmu pre rozpoznávanie vzoru nad vyskladanou maticou

Algoritmus pre rozpoznávanie vzoru bude postavený na metóde nazývanej zhluk. Úlohou algoritmu bude nájsť smer príchodu, zenitový a axiálny uhol spŕšky s určitou presnosťou pre jej neskoršiu rekonštrukciu a určenie možného zdroja spŕšok. Ako už bolo vyššie spomenuté, finálne rozpoznávanie sa bude diať už nad vyskladanou maticou pomocou modifikovanej sumačnej metódy, pozri obrázok 7–4. Samotná spŕška je v matici reprezentovaná ako počet fotoelektrónov. Počty sú uchované pomocou celočíselnej hodnoty v jednotlivých pixeloch vyskladanej matice. Hodnoty väčšie ako hranica šumu pozadia môžu reprezentovať spŕšku. Algoritmus sa bude zameriavať najmä na tieto hodnoty pixelov a na základe ich pozície v rámci vyskladanej matice vykoná rekonštrukciu smeru príchodu spŕšky.

7.2.1 Návrh vzoru spŕšky pre rozpoznávanie

Pri rekonštrukcii smeru príchodu spŕšky si potrebujeme zdefinovať jej vzor, na základe ktorého bude jednoznačne identifikovaná. Vzor je všeobecná definícia spŕšky potrebná pre rozpoznávanie pomocou metódy zhluk. Spŕšku kozmického žiarenia nám predstavuje priamočiara neprerušovaná línia. Neprerušovaná v tom zmysle, že hodnoty pixelov, ktoré reprezentujú spŕšku nebudú menšie ako je požadovaná hranica oddeľujúca šum pozadia od reálneho výskytu spŕšky. Priamočiara a neprerušovaná hovorí, že hodnoty pixelov pozdĺž celej línie budú vždy väčšie ako je pozadie šumu. Dobré zdefinovanie vzoru spŕšky je jednou z najdôležitejších častí pri rozpoznávaní jej smeru príchodu.

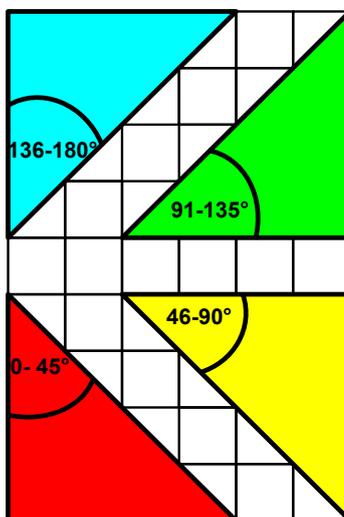
7.2.2 Návrh spôsobu prechádzania vyskladanou maticou

Kedže sa nevie, ktorým smerom prišla sprška a taktiež sa nevie, kedy ju detektor zachytil (v akej výške), je potrebné vymyslieť algoritmus, ktorý zanalyzuje celú maticu tak, aby zobral do úvahy každú možnosť. Vie sa, že matica má rozmery 36x36 pixelov. Každá sprška bude mať začiatkový bod, ktorý môže byť reprezentovaný ľubovoľným pixelom z matice a bude obsahovať aj koncový bod, taktiež reprezentovaný pixelom z matice rôznym od začiatového pixelu. To môžeme aplikovať na celú maticu pomocou cyklov.

Pre každý bod matice sa postupne prejdú všetky ostatné body v matici, ktoré nie sú na tej istej pozícii ako začiatkový bod a budú sa považovať za koncové body. Takto sa prejde každá možnosť zachytenia spršky detektorom JEM-EUSO. Sprška môže byť predstavená ako čiara medzi začiatkovým a koncovým bodom. Pixely, ktoré sú v prieniku s priamkou musia obsahovať hodnoty väčšie ako je hranica šumu. Ak tomu tak nie je a hodnota ľubovoľného pixelu je menšia alebo rovná šumu pozadia, tak medzi začiatkovým a koncovým bodom sa nenachádza sprška, pretože nespĺňa podmienky spojitosti vzoru spršky, pozri sekciu 7.2.1.

Pre určenie bodov, ktoré sú v prieniku s priamkou sa použije algoritmus, ktorý je založený na pravouhlých trojuholníkoch. Základným princípom je, že línia medzi prvým a posledným bodom nám predstavuje preponu pravouhlého trojuholníka. Uhol je reprezentovaný medzi zvislou osou matice a preponou trojuholníka. Maticu si rozdelíme do štyroch kvadrantov podľa uhla spršky. Potom na základe uhla sa pre výpočet nasledujúceho bodu spršky, nachádzajúcim sa medzi začiatkovým a koncovým bodom použije ten pravouhlý trojuholník, v ktorom kvadrante sa priamka nachádza, čo je vidieť aj na obrázku 7–6.

Pri kvadrante zastrešujúcom rozmedzie uhlov medzi 0 až 45 stupňov sa nasledujúci pixel vypočíta tak, že na zvislej osi sa zvýši hodnota pozície v matici o 1, čiže ak aktuálna pozícia v matici na zvislej osi je 5 po tomto kroku bude 6. Pre výpočet



Obr. 7–6 Trojuholníky reprezentujúce kvadranty, ktoré sa použijú pri výpočte nasledujúceho bodu na základe uhla spŕšky.

pozície na vodorovnej osi je už uhol a aj veľkosť protiláhlej pripony známa. Analogicky rovnaký výpočet platí aj pre tretí kvadrant, na obrázku 7–6 označený zelenou farbou, len na zvislej osi sa hodnota 1 odpočítava. Pri uhle v rozmedí 46-90 stupňov sa nasledujúci pixel určí podobne ako v prvom kvadrante no s tým rozdielom, že hodnota o jeden sa zväčší na vodorovnej osi a pomocou vzorca sa dopočítava hodnota pozície na zvislej osi. Analogicky to platí aj pre posledný, štvrtý kvadrant.

7.3 Návrh algoritmu pre integráciu do ESAF rámca

Rekonštrukcia a rozpoznávanie spŕšok v rámci ESAF sú zahrnuté v Rekonštrukčnom balíku. Vďaka jeho návrhu, pozri sekciu 6.2, je možné vytvárať a pridávať nové moduly. Na základe sekcie 6.2.1 musí modul dediť od triedy RecoModule. Taktiež sa pre analýzu musí zimplementovať vlastná logika v metóde pre spracovanie spŕšky, ktorá je dodávaná ako parameter. Podľa štruktúry reprezentujúcej zachytenie dát pre šum a spŕšku na povrchu detektora v rámci ESAF, viď sekciu 6.2.2 a podľa štruktúry dát, nad ktorou sa vykonáva rozpoznávanie pomocou metódy zhľuku,

pozri návrh v sekcii 7.2.1 a následnú implementáciu, viď sekcii 8.1, je vidieť, že sú nekonzistentné. Preto je potrebné vytvoriť konvertér z dát, ktoré produkuje ESAF na dáta, ktoré je možné analyzovať metódou zhľuku.

Vie sa, že pozícia pixelov na povrchu detekčného modulu fotónov je v rámci ESAF reprezentovaná pomocou milimetrov. Konkrétne hodnoty pre konkrétny pixel sa získajú na základe identifikačného čísla pixelu. Na opačnej strane pozícia pixelu na detekčnom module v rámci dát, ktoré sú analyzované metódou zhľuku je reprezentovaná ako hodnota indexov v dvojrozmernom poli, reprezentujúceho maticu. Konverziu je možné postaviť použitím identifikačného čísla pixelu, pretože pomocou neho sa získa pozícia na detektore, ktorá je pre identifikačné číslo stále rovnaká.

Povrch detekčného modulu fotónov v rámci ESAF je rozdelený na menšie štvorcové časti nazývané aj elementárna bunka. Ich rozloženie na základe hodnôt identifikačných čísel je následovné. Prvá elementárna bunka, tá ktorej hodnoty pixelov sú najmenšie, sa nachádza v ľavom dolnom rohu povrchu. Posledná elementárna bunka sa nachádza v pravom hornom rohu. Po povrchu sú rozložené systémom, po riadkoch od spodu smerom hore, kde v jednotlivých riadkoch sú rozložené zľava doprava. Jedna elementárna častica obsahuje 256 pixelov a na danom module sa nachádza deväť elementárnych buniek.

Každá elementárna častica sa skladá zo štyroch menších fotónových multiplikátorov. Každý multiplikátor obsahuje 64 pixelov. V rámci elementárnej bunky sú multiplikátory rozdelené tak, že multiplikátor, ktorého identifikačné číslo pixelov sú najmenšie, t.j. od 1 do 64 sa nachádza vpravo dole elementárnej bunky. Potom ďalší fotónový multiplikátor, kde čísla pixelov sú v rozmedzí 65 až 128 sa nachádza vpravo hore. Tretí multiplikátor s pixelmi 129 až po 192 sa nachádza vľavo dole a posledný s pixelmi od 193 po 256 sa nachádza v ľavom hornom rohu elementárnej bunky.

Ako posledné sa určí rozloženie pixelov v rámci fotónového multiplikátora. Prvý pixel, ktorého hodnota je jeden sa nachádza v ľavom hornom rohu. Posledný pixel s

identifikačným číslom 64 sa nachádza v pravom dolnom rohu. Rozloženie pixelov je určené nasledujúcou štruktúrou. Hodnoty narastajú postupne v rámci stĺpca smerom zhora dole. Kde ak sa narazí na spodnú hranu fotónového multiplikátora, tak sa pokračuje od vrchu o jeden stĺpec v pravo.

Na základe vyššie uvedených faktov ohľadom pozície pixelu určených pomocou ich identifikačného čísla je rozloženie pixelov v rámci jednotlivých úrovni pravidelné, a preto je možné napísať algoritmus, ktorý skonvertuje na základe identifikačného čísla pixelu jeho pozíciu zo štruktúry v rámci ESAF na štruktúru používanú metódou zhuk. Po tejto konverzii dát je možné aplikovať metódu zhukovania na dáta generované rámcom ESAF.

7.4 Návrh pre implementáciu Houghovho priestoru

Houghova metóda prechádza každým pixelom vyskladanej matice. Pre tento pixel na základe analýzy, pozri sekciu 5.6, sa určí kolmá vzdialenosť priamok k počiatku súradnicovej sústavy pre všetky uhly. Tieto dáta sa uložia do poľa a vznikne tak pole informácií o všetkých pixeloch a ich kolmej vzdialenosti pre všetky smery. Pole informácií sa nazýva Houghov priestor a obsahuje informáciu o vzoroch skúmanej matice.

Pri zoskupovaní bodov, ktoré tvoria hľadaný vzor vyberáme tie body, ktorých kolmá vzdialenosť uchovaná v Houghovom priestore je podobná, kde maximálna odchylka je $r \pm dr$ kde $dr = (\text{dĺžka hrany pixelu}/2)$ na základe analýzy, vid' sekciu 5.6. Takto získané body predstavujú vzor.

Keďže Houghova transformácia vyhľadáva body pozdĺž línie, kde neberie do úvahy ich hodnoty, musí sa následne splniť podmienka spojitosti. Podmienka spojitosti sa aplikuje tak, že pre zoradené hodnoty pixelov nachádzajúce sa pozdĺž línie zistí, či vzdialenosť medzi stredmi pixelu je menšia alebo rovná $\sqrt{2}$.

8 Implementácia algoritmov na rozpoznávanie vzoru

Na začiatku sa pre overenie správnosti distribúcie šumu použil jazyk IDL. Neskôr sa kvôli efektívnosti a aj lepšej integrácii do rámca ESAF implementácia algoritmov vykonávala pomocou jazyka C++. Zimplementuje sa jednoduchá verzia Hougovej transformácie a taktiež aj metóda zhlukovania. Potom sa zintegruje metóda zhlukovania do rámca ESAF. Následne sa porovnajú výsledky medzi Hougovou transformáciou a metódou zhlukovania pri použití rôznych parametrov definície spŕšky.

8.1 Implementácia rozpoznávania pomocou metódy zhlukovania

Vstupné dáta simulácie sa nachádzajú v textovom súbore. Každý pixel s jeho pozíciou, GTU a hodnotou šumu je zapísaný na jednom riadku. Ako prvé sa preto vytvoril parser, ktorý načítal dáta zo súboru do pamäte. Na základe návrhu algoritmu sa zimplementuje metóda pre vyskladanie matice, pozri kapitolu 7. Je to prvý krok po načítaní dát. V tomto kroku sa vygenerovaná udalosť pomocou upraveného sumačného algoritmu, viď sekciu 7.1.2, prekonvertuje a výsledok sa uloží do novej štruktúry pre zosumovanú udalosť, zatiaľ čo stará štruktúra udalosti sa uvoľní z pamäte.

Po upravení sumačnou metódou sú dáta pripravené na použitie pre algoritmus na rozpoznávanie vzorov. Metóda postupne prechádza každou udalosťou. Aby sa minimalizoval výpočtový čas, kontroluje sa, či hodnoty začiatkových a koncových bodov sú väčšie ako je požadovaná hranica oddeľujúca šum od signálu. Ak áno, tak pomocou pravouhlých trojuholníkov, pozri sekciu 7.2.2, sa postupne od prvého bodu zistí pozícia nasledujúceho pixelu a skontroluje sa, či jeho hodnota je väčšia ako hranica pozadia. Ak je, pokračuje sa hľadaním ďalšieho pixelu až pokiaľ sa nedôjde ku koncovému pixelu. Potom sa nájdená spŕška uloží do štruktúry pre nájdené spŕšky. Ak

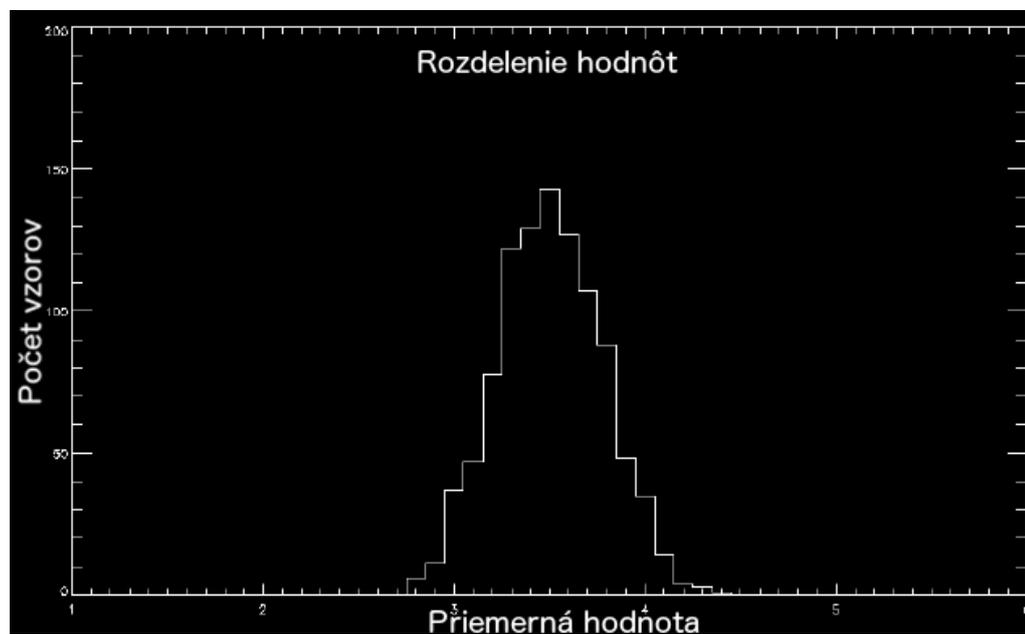
nie, tak sa ukončí hľadanie aktuálnej spŕšky a pokračuje sa v hľadaní ďalšej spŕšky.

Po ukončení rozpoznávacieho algoritmu sa spustí vyhodnocovací algoritmus. Jeho úlohou je vyhodnocovanie spŕšok a generovanie výsledkov, o ktoré sa budú opierať výsledky práce.

8.2 Implementácia správnosti použitia metódy zhlukovania

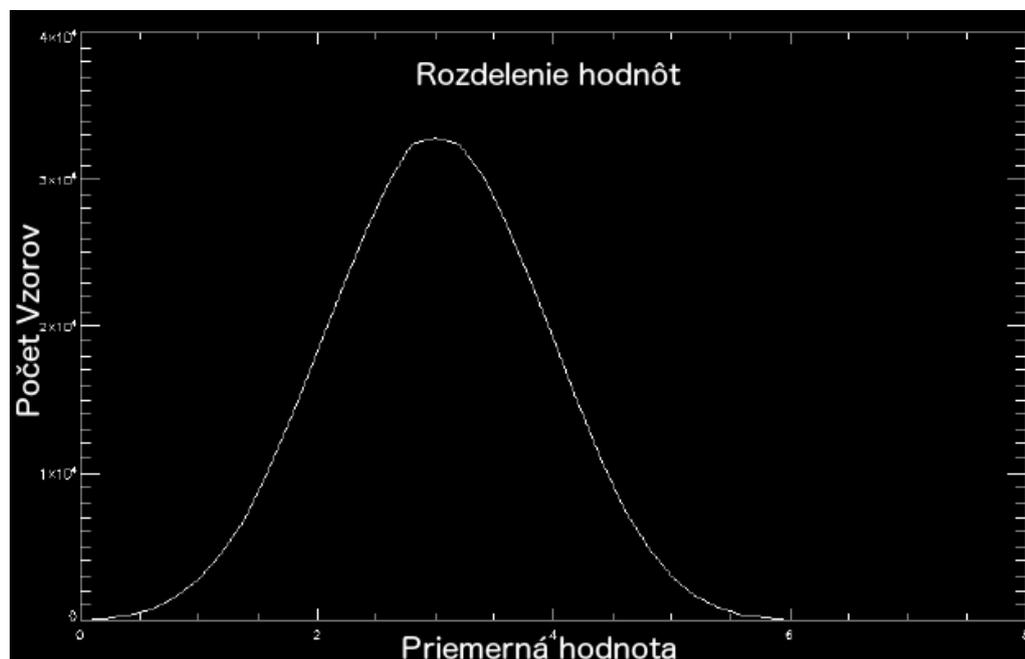
Pre otestovanie správnosti použitia metódy zhlukovania na algoritmus pre rozpoznávanie spŕšok sa v rámci návrhu porovná jednoduchá úloha, kde sa porovnávajú výsledky medzi teóriou a testovacím riešením pomocou zhlukovania. Na riešenie jednotlivých problémov sa používal programovací jazyk IDL (Interactive Data Language). Pomocou tohto jazyka boli algoritmicky riešené prvé úlohy, kde sa generovali matice 8×8 s hodnotami jednotlivých pixelov 0,1. Porovnávalo sa riešenie predpovedané teóriou s výsledkami získanými algoritmom hľadajúcim viaceré vzory v stĺpcoch matice. Cieľom bolo dokázať zhodnosť výsledkov, ktoré sme dosiahli pomocou teórie a pomocou náhodne vygenerovaných matíc, ktoré sme generovali pomocou daného programu. Opis programu je možné vidieť v sekcii 8.3. Neskôr jednotlivé pixely matice obsahovali hodnoty v rozmedzí 0 až 7, ktoré boli uniformne generované. Uniformnosť sme si dokazovali tak, že sme získali priemernú hodnotu pixelu pre všetky matice. Priemernú hodnotu pixelu matice sme získali predelením súčtu hodnôt pixelov ich počtom. Na základe početnosti priemerných hodnôt pixelov sme si výsledok zobrazili ako histogram, ktorý je vidieť na obrázku 8–1, ktorý nám dokazuje uniformnosť generovaných matíc. Toto všetko bolo realizované pomocou jazyka IDL.

Keďže výkonnosť tohto jazyka je neporovnateľne menšia ako jazyka C++, tak sme pre výpočty začali používať jazyk C++. Po dokázaní uniformnosti generovaných matíc sme začali hľadať rôzne vzory pre rôzne dĺžky a rôzne smery (uhly). Po vygenerovaní požadovaného množstva matíc sa pre každý pixel každej matice hľadajú vzory rôznych dĺžok pre rôzne smery. Zatiaľ sa do úvahy berie každý platný vzor



Obr. 8–1 Zobrazenie histogramu priemernej hodnoty pixelu matice.

v rámci matice bez ohľadu na hodnotu v danom pixely. Pre každý nájdený vzor sa uchováva hodnoty o jeho dĺžke, uhle, pod ktorým sa našiel, pixel odkiaľ vzor začína a taktiež priemerná hodnota vzoru. Pre kontrolu sa môže zobraziť početnosť vzorov s priemernou hodnotou pixelu buď pre všetky dĺžky a smery, alebo sa môže vybrať na základe dĺžky vzoru a smeru, pod ktorým bol nájdený. Pre ilustráciu výsledku sa vybralo zobrazenie závislosti počtu nájdených spířšok od priemernej hodnoty patternu a jeho dĺžky. Graf na obrázku 8–2 zobrazuje danú závislosť pre vzory dĺžky 5 a pre všetky smery (uhly).



Obr. 8–2 Zobrazenie závislosti počtu nájdených spířšok od priemernej hodnoty patternu a jeho dĺžky (5).

8.3 Testovanie pomocou metódy zhlukovania

Pre otestovanie algoritmu vytvoreného na základe metódy zhlukovania sa zadefinuje jednoduchá úloha. Hľadanie vzorov sa bude vykonávať na maticiach o rozmeroch 8×8 , kde každý pixel bude predstavovať hodnotu 0 alebo 1, ktoré budú generované uniformne programom. Takýchto matíc sa vygeneruje 10 000.

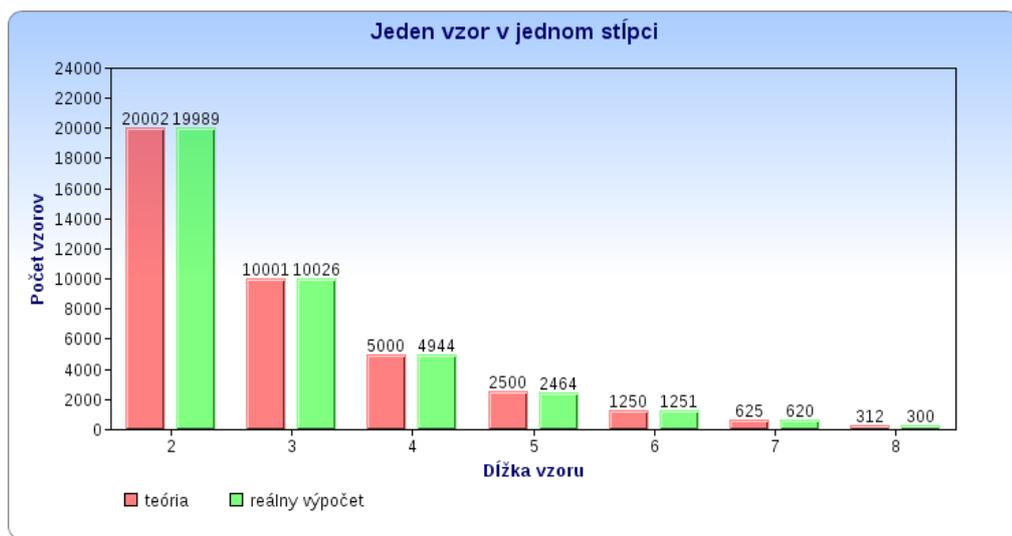
Nad týmito dátami sa vykoná jednoduché rozpoznávanie metódou zhlukovania pre vertikálne vzory. Pre zjednodušenie sa hľadajú len vertikálne vzory pre jediný uhol 90 stupňov. Na porovnanie výsledkov, ktoré sa dosiahnú spustením samotného programu sa vypočítajú aj teoretické pravdepodobnosti a na konci sa porovnajú. Pravdepodobnosť sa bude počítať pre dva prípady. Pre jednotlivé vzory je pravdepodobnosť P nájdenia vzoru o dĺžke DP pri matici s dĺžkou strany DS nasledovná, vid' sekciu 8.3.1 a sekciu 8.3.2.

8.3.1 Jediný vzor v jednom stĺpci

Vzor sa môže v danom stĺpci nachádzať len raz, čo znamená, že ak sme v danom stĺpci našli vzor o dĺžke 2, už sa viac vzor s takouto dĺžkou v danom stĺpci nehľadá. Výpočet pravdepodobnosti bude, kde počet možných hodnôt pixelu PP je rovné dvom (pixel môže nadobúdať hodnotu 0 alebo 1)

$$P = (1/PP)^D P * DS$$

Po vypočítaní teoretických pravdepodobností a náhodnom spustení programu sa dosiahnú výsledky, ktoré sú znázornené na obrázku 8–3.



Obr. 8–3 Zobrazenie teoretickej pravdepodobnosti početnosti a reálneho výsledku.

8.3.2 Viacero vzorov v jednom stĺpci

Vzor sa môže v danom stĺpci nachádzať viackrát, čo znamená, že ak sa v danom stĺpci nájde vzor o dĺžke 2, tak v danom stĺpci sa stále hľadá vzor o dĺžke dva, no nemôže to byť ten istý vzor a taktiež musí byť medzi vzormi aspoň jeden nulový pixel, ktorý tieto dané vzory oddelí. Potom výpočet pravdepodobnosti bude, kde

počet možných hodnôt pixelu PP je rovné dvom (pixel môže nadobúdať hodnotu 0 alebo 1)

$$P = (1/PP)^D P * (DS * (DS - DP + 1))$$

Po vypočítaní teoretických pravdepodobností a náhodnom spustení programu sa dosiahnú výsledky, ktoré sú znázornené na obrázku 8–4



Obr. 8–4 Zobrazenie teoretickej pravdepodobnosti početnosti a reálneho výsledku.

Na tomto jednoduchom príklade je vidieť, že pomocou metódy zhlukovania sa dosiahnú takmer zhodné výsledky s teoretickou pravdepodobnosťou, čo je vidieť na obrázkoch 8–3 a 8–4.

8.4 Implementácia modulu do rámca ESAF

Ako implementačný jazyk sa použil C++, pretože rámec ESAF je celý napísaný v tomto jazyku. Po získaní prístupu k zdrojovým kódom rámca ESAF sa vytvoril základný prázdny modul pre rekonštrukčný balíček na základe sekcie 7.3. Zaregistroval sa do rámca a spustil sa kód pre rekonštrukciu na otestovanie, či bol modul správne naprogramovaný. Po tomto kroku sa vložili do rámca zdrojové kódy založené na rozpoznávaní pomocou zhluku, viď sekciu 8.1. V tomto kroku už bolo všetko pripravené na implementáciu konvertéra zo štruktúry spíšky na maticovú štruktúru

používanú metódou zhlukovania.

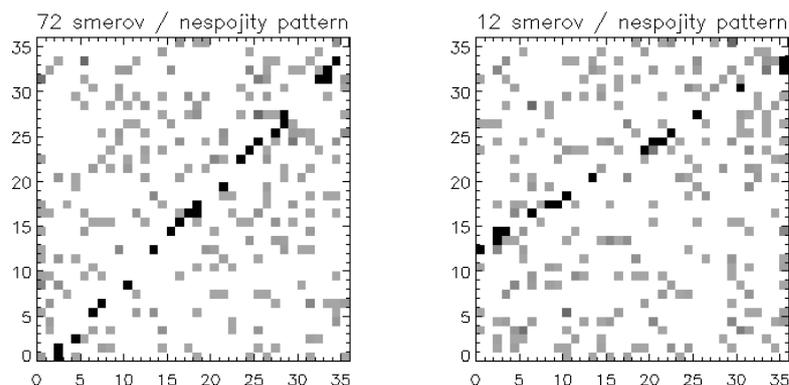
Podľa návrhu sekcie 7.3 sa vie, že každý pixel má svoje jednoznačné číslo, ktoré je v jednotlivých vrstvách rozložené pravidelne. Podstatou algoritmu je, že vždy sa ako prvé určila pozícia pixelu v rámci fotónového multiplikátora, ktorý sa už predstavoval ako matica. Následne sa pomocou čísla pixelu zistilo, v ktorom multiplikátore v rámci elementárnej častice sa nachádza a pripočítal sa offset pre súradnice pixelu. Po tomto kroku sa pripočítal offset, ktorý sa určil zistením pozície elementárnej častice v detekčnom module fotónov.

Týmto algoritmom sa prekonvertované pixely ukladali už do maticovej štruktúry používanej rozpoznávacou metódou zhlukovania. V rámci modulu sa spustili analyzačné kódy zhlukovacej metódy.

8.5 Implementácia Houghovho priestoru a nespojitosť vzoru

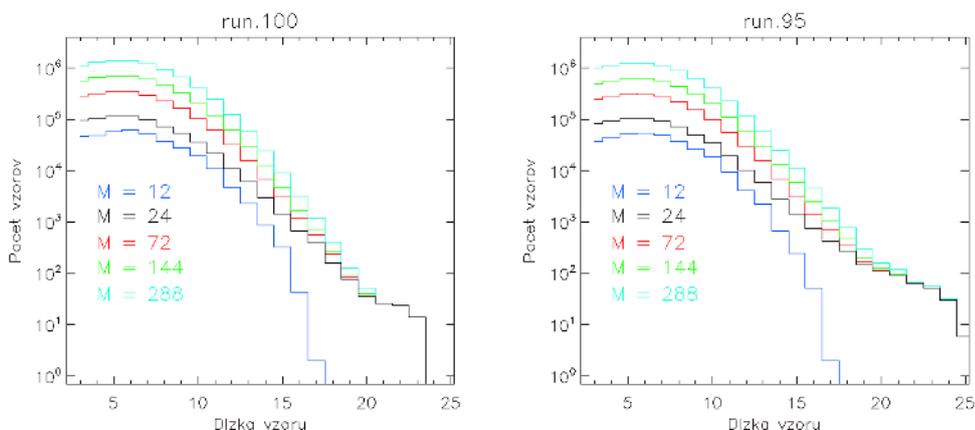
Na základe návrhu Houghovej transformácie, pozri sekciu 7.4, bol vytvorený program, ktorým boli vykonané nasledujúce výpočty. Na obrázku 8–5. sú zobrazené dve matice 36x36 pixelov (PDM), v ktorých sa nachádzali najdlhšie vzory v prehľadávaných maticiach z jedného behu spúšťačej štatistiky. To znamená matice, ktoré prešli spúšťačom z nagenеровanej miliardy matíc, ktorá znamená meranie v dĺžke 2500 sekúnd na jednom PDM (keďže GTU je 2,5 μ s). Simulácie pozadia sa vykonávali v behoch po 109 PDM-iek. Takýchto behov bolo vykonaných pár stoviek pričom len približne každý desiaty vyprodukoval prespúšťačie udalosti. Pre overenie výsledkov analýzy Houghovou metódou sa využili výsledky behu 100 (run.100, pokiaľ nie je v ďalšom texte tejto kapitoly výslovne uvedené inak, ide o výsledky z tohto behu). Z 31 matíc každej udalosti (z behu 100), ktorá prešla spúšťačom bola vytvorená jedna matica (modifikovaná sumočná metóda), na ktorej sa vzory hľadali Houghovou metódou. Na obrázku 8–5. vpravo je zobrazená matica, na ktorej s použitím Houghovej metódy pre $M=12$, tj. pre $\alpha = 30^\circ$ bol nájdený najdlhší vzor pozostávajúci zo

17 pixelov. Pixely nájdeného najdlhšieho vzoru sú zvýraznené čiernou farbou. Pre $M=72$, tj. pre $\alpha = 5^\circ$ bol nájdený najdlhší vzor pozostávajúci z 23 pixelov (ľavý panel obrázku).



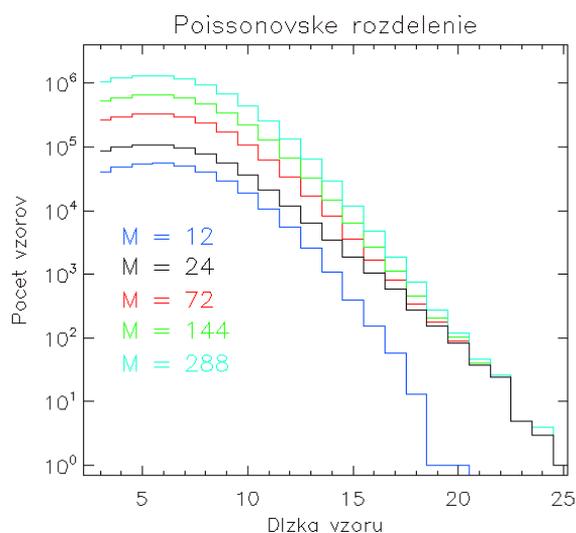
Obr. 8–5 Najdlhšie nespojité vzory pre $M=72$ (ľavý panel obrázku) a $M=12$ (pravý panel).

Početnosť nájdených vzorov ako funkciu dĺžky vzoru v prípade nespojitého vzoru ukazuje obrázok 8–6. pre $M=12$ (modrý histogram), $M=24$ (čierny histogram), $M=72$ (červený histogram) a $M=144$ (zelený histogram). Početnosť rastie s nárastom počtu smerov. Od určitého momentu rastie preto, lebo jedným pixelom prechádza viacero smerov. Pre zobrazenie vplyvu získanej štatistiky simulácie na obrázku 8–6. sa uvádzajú aj výsledky z ďalšieho behu (beh 95).



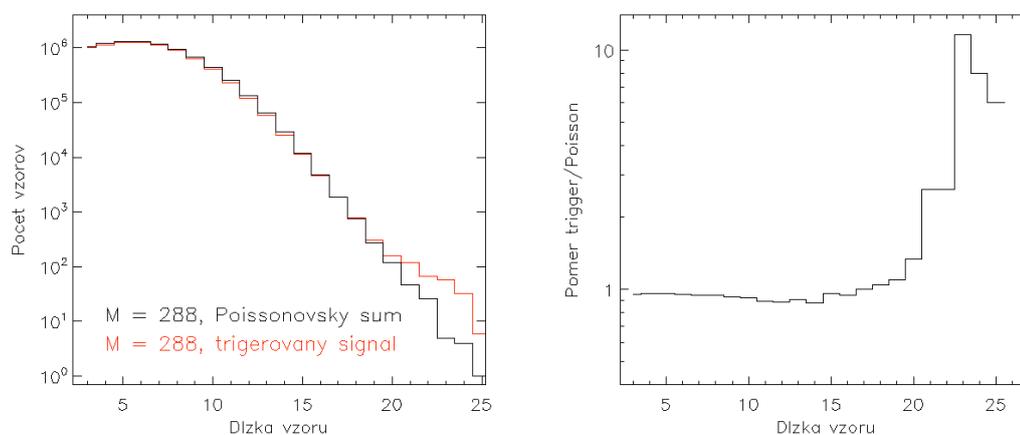
Obr. 8–6 Závislosť počtu nespojitých vzorov od dĺžky vzoru pre rôzne počty smerov v Houghovej metóde (viď. popis v texte) pre signál po prechode spúšťacou schémou.

Pri pohľade na obrázok 8–6. je viditeľné navýšenie počtu nájdených vzorov pre väčšie dĺžky mimo trendu kriviek na obrázku. Inak povedané, našlo sa viac dlhých vzorov než by sa podľa trendu kriviek očakávalo. Toto je spôsobené spúšťacím algoritmom, ktorý vyhľadáva matice, kde pozdĺž priamych línií je suma pixelov väčšia než prah druhej spúšťacej úrovne. Ten je 143, a teda pri dĺžkach vzoru ~ 25 pixelov je stredná hodnota týchto pixelov 5. To je v súlade s vyžadovanými vlastnosťami pixelov v použitej metóde skladania matíc. Navýšenie pre dlhé vzory je spôsobené spúšťacím algoritmom. Aby sa tento záver dokázal a ilustroval odlišnosť od situácie pre nespúšťací signál, vykonali sa analýzy na maticiach len s Poissonovským šumom, bez ich filtrovania spúšťacou schémou. Výsledkom je krivka bez navýšenia v oblasti dlhších patternov.



Obr. 8–7 Závislosť počtu nespojitých vzorov od dĺžky vzoru pre rôzne počty smerov v Houghovej metóde (vid'. popis v texte) pre Poissonovský signál bez prechodu spúšťacom.

Na obrázku 8–7 je uvedená závislosť počtu nájdených vzorov od dĺžky vzorov pre nespúšťací signál. Porovnanie rozpoznávania vzorov pre 288 smerov je uvedené na obrázku 8–8. Na pravom paneli obrázku je podiel medzi početnosťami nájdených vzorov predelený početnosťou signálu v nespúšťacom signále. V spúšťacom signále sa dlhé vzory nachádzajú až desaťkrát častejšie ako v nespúšťacom.



Obr. 8–8 Porovnanie počtu vzorov pre udalosti vyselektované spúšťačom a generovaný signál(ľavý panel) a ich pomer.

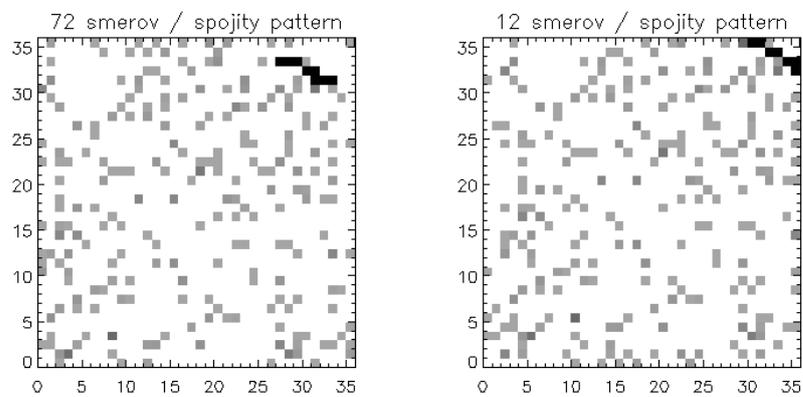
8.6 Implementácia Houghovho priestoru a spojitosť vzoru

Z hľadiska rozpoznávania spršok kozmického žiarenia je potrebné zaviesť podmienku spojitosti vzoru. Sprška generuje signál, ktorý nepreskočí žiaden pixel. Možným vplyvom nespojitosti sa krátko zaoberá časť práce Vplyv nespojitosti vzoru v ďalšom texte diplomovky.

Spojitosť signálu v Houghovej metóde sa môže zadefinovať maximálnou povolenou vzdialenosťou stredov pixelov v nájdenom vzore. Stredy dvoch pixelov spojitého vzoru majú maximálnu vzdialenosť ak sú ich stredy spojené ich diagonálami. Keď ich spojnice zvierajú so stranou pixelu 45° uhol. Vtedy sú stredy pixelov vzdialené $\sqrt{2}$ (jednotkou je dĺžka strany pixelu). Podmienka spojitosti vzoru tak môže byť definovaná tak, že vzor je spojitý keď každý z jeho pixelov je od stredov susedných pixelov vzdialený maximálne o hodnotu $\sqrt{2}$. Ak je vzdialenosť stredov pixelov väčšia, tak je vzor nespojitý.

Po aplikovaní podmienky spojitosti pre $M=12$ a $M=72$ pri $dr=0.5$ pixelov sa dĺžka nájdených vzorov zmenšila. Na obrázku 8–9. sú znázornené príklady najdlhších nájdených vzorov. Ľavý panel obrázku ukazuje vzor s maximálnou nájdenou dĺžkou 7 pixelov pre $M=12$ t.j. pre $\alpha = 30^\circ$. Pravý panel obrázku ukazuje najdlhší vzor

nájdenny pre $M=72$, tj. pre $\alpha = 5^\circ$ dlhý 8 pixelov.



Obr. 8–9 Najdlhšie spojité vzory pre $M=72$ (ľavý panel obrázku) a $M=12$ (pravý panel).

9 Overovanie výsledkov

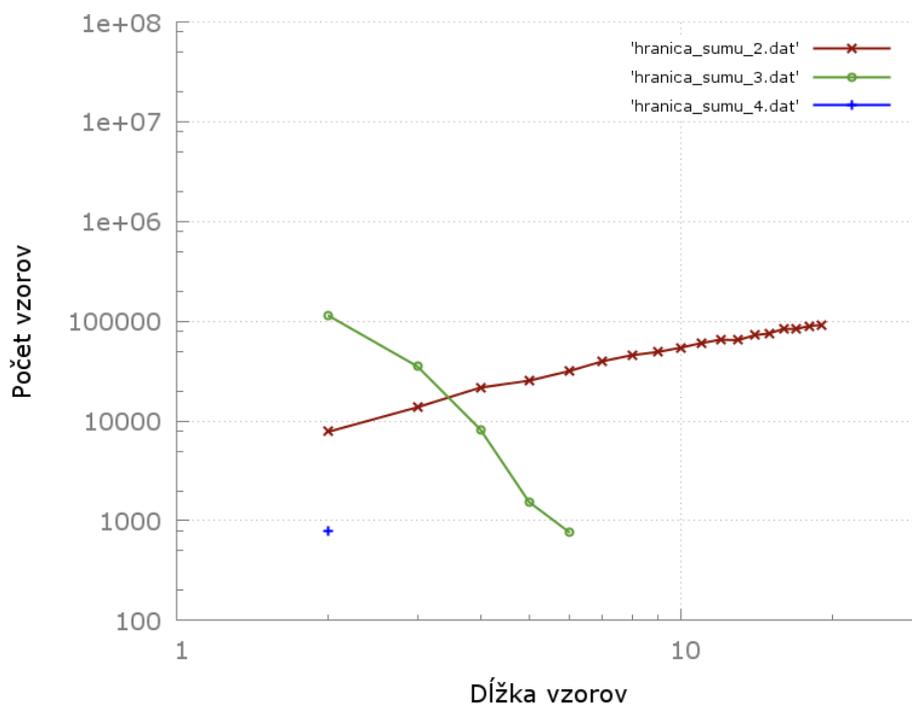
V rámci JEM-EUSO projektu sa vygenerovali behy, ktoré obsahovali len šum pozadia. Z týchto behov sa vybrali tie udalosti, ktoré prešli prvou a druhou spúšťacou úrovňou, čo znamená, za náhodným Poissonovským generovaním pozadia vznikol útvar podobný sprške. Táto kapitola je zameraná na zistenie vplyvu hranice oddeľujúcej šum od pozadia na výsledky a či Poissonovské rozdelenie šumu môže vytvoriť spršku, ktorá by bola podobná reálnej sprške.

9.1 Overenie vplyvu závislosti hranice šumu na výsledky

Jednou z definícií spršky je jej spojitosť, kde hodnota počtu fotónov zachytených na pixeli detektora nesmie byť menšia ako je stanovená hranica. Taktiež pixely s hodnotou väčšiou ako je stanovená hranica sa musia nachádzať vedľa seba, pozri sekciu 7.2.1. Po aplikovaní modifikovanej sumačnej metódy, vid' sekciu 7.1.2, obsahuje matica pixely s hodnotou 2 a viac alebo 0. Preto sa pri rozpoznávanom algoritme určí viacero hraníc určujúcich spojitosť spršky a porovnávajú sa ich výsledky. Ako hodnotiacia metodika sa použije závislosť dĺžky spršok od ich počtu. Na vyhodnocovanie sa použil beh s číslom 100. Závislosť početnosti na základe hranice je vidieť na obrázku 9–1.

Na obrázku 9–1 predstavuje červená farba početnosť spršok pre hranicu s hodnotou 2 už po aplikovaní modifikovanej sumačnej metódy. Krivka má stúpajúcu tendenciu, pretože algoritmus vynecháva spršky, ktoré sú podmnožinou inej väčšej spršky. Tento graf naznačuje doslovné zahĺtenie spršok v signále. To znamená, že hranica s hodnotou 2 je menšia alebo rovná priemernej hodnote šumu, a preto algoritmus považuje za spršku aj samotný šum. Takáto hodnota hranice je nepoužiteľná pri overovaní výsledkov.

Početnosť spršok, kde hranicu reprezentuje hodnota 3 je na obrázku 9–1 znázornená



Obr. 9–1 Závislosť počtu vzorov od ich dĺžky pre rôzne hranice šumu.

zelenou farbou. Krivka ma klesajúcu tendenciu. Počet kratších spŕšok je vždy väčší ako tých dlhších. Hranica spŕšky je tu väčšia ako je priemerná hodnota pozadia, čím sa zaručilo nezahľtenie spŕšok s pozadím a zároveň nie je veľmi vysoká, čo má za následok dodržanie spojitosti vzoru aj pri dlhších vzoroch.

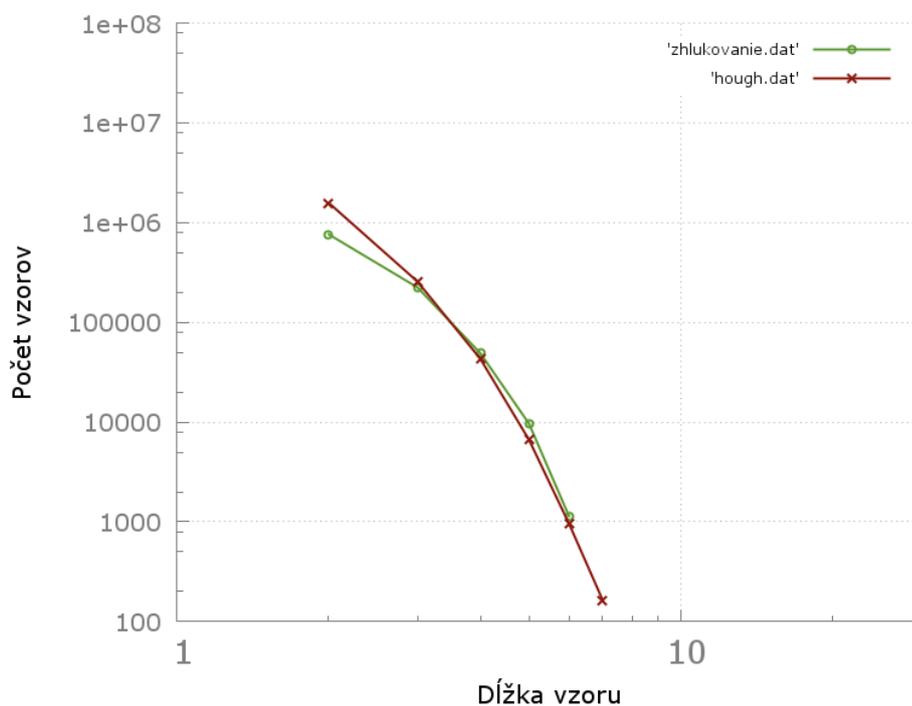
Počet spŕšok voči ich dĺžke pri hranici s hodnotou 4 je na obrázku zobrazená modrou farbou. S touto hranicou sa nenašli spŕšky s dĺžkou väčšou ako 2. Hodnota hranice je už o dosť väčšia ako je priemerná hodnota šumu s Poissonovským rozdelením, a preto je podmienka spojitosti vzoru splnená len pri najkratších možných spŕškach.

Z obrázka 9–1 vyplýva, že početnosť nájdených spŕšok silne závisí na definícii vzoru spŕšky.

9.2 Vyhľadávanie spršok v Poissonovom šume

Pre zistenie či je možné zameniť spršku, ktorá vznikla náhodne v šume s reálne zachytenou sprškou sa zistí závislosť dĺžky spršky od jej početnosti v rámci behov, kde bolo generované pozadie šumu. Do analýzy sa zahrnie dokopy desať behov. Cieľom je rozpoznať čo najkratšie vzory, čo by znamenalo, že sprška priletela pod veľkým (ostrým) uhlom a jej rekonštrukcia by bola veľmi náročná.

Na analýzu behov sa použili dva algoritmi. Metóda zhlukovania 8.1 a Houghova metóda 8.5. Obe metódy vyhľadávajú potencionálne spršky nad udalosťami, ktoré boli vyskladané upravenou sumačnou metódou 7.1.2. Ako hranica šumu sa použije hodnota 3 na základe vplyvu závislosti hranice šumu na výsledky, viď sekciu 9.1. Výsledky metód sú znázornené na obrázku 9–2.



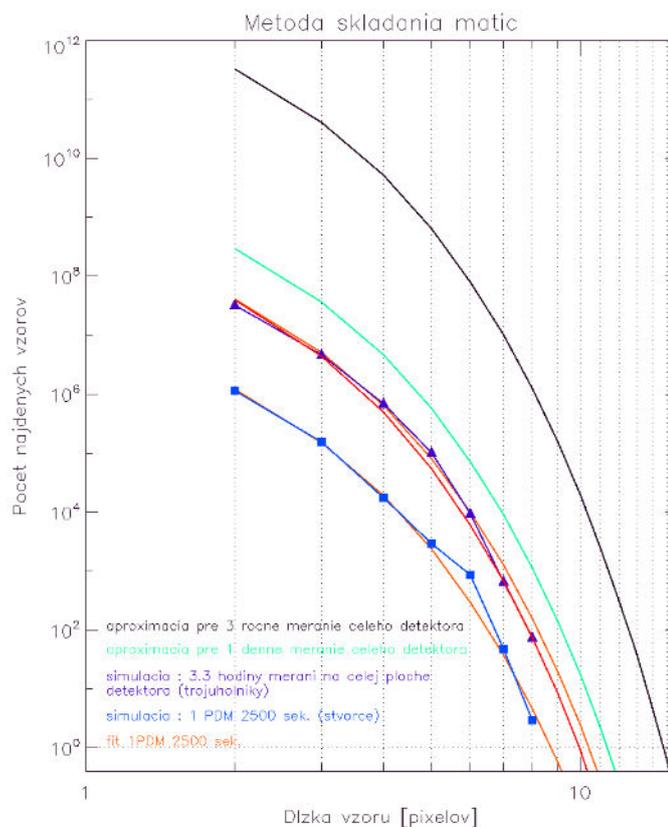
Obr. 9–2 Závislosť počtu vzorov od ich dĺžky získaných Houghovou metódou a metódou zhlukovania.

Výsledky metódy zhlukovania sú na obrázku 9–2 znázornené zelenou farbou. Výsled-

ky metódy Houghoveho priestoru sú zobrazené červenou farbou. Ako je vidieť, výsledky dvoch rôznych algoritmov sú kvalitatívne zhodné. Najdlhší vzor v Poissonovsky rozdelenom šume sa podarilo nájsť o dĺžke 7, na dátach zodpovedajúcich rádovo hodinám merania celého JEM-EUSO detektora. Štatisticky je možné odhadnúť najdlhší spojitý falošný vzor pre trojročné meranie JEM-EUSO detektora. Jeho dĺžka nepresiahne 15 pixelov a teda dĺžku 7,5 kilometra v priemete na zemskom povrchu. Keďže prvý bod spŕšky zaznamenaný detektorom sa nachádza vo výške približne 13-15 kilometrov, zenitový uhol takejto udalosti bude 29,98 stupňa pre 13 kilometrovú výšku a 26,56 stupňa. To znamená, že jeho zenitový uhol príchodu je menší ako je uhol najkratšej rekonštruovateľnej spŕšky v rámci JEM-EUSO projektu a tým pádom nehrozí pomýlenie si falošnej spŕšky s reálnou.

9.3 Porovnanie výsledkov s vyskladávaním Houghovou transformáciou

Zistením závislosti početnosti od dĺžky sa zaoberal vo svojej diplomovej práci aj Staroň [35]. Na vyskladávanie udalosti použil vlastnú metódu založenú na Houghovej transformácii. Jeho rozpoznávanie bolo vykonané nad upravenými údajmi rozličnou metódou, aká sa použila v rámci tejto diplomovej práce. Výsledky z rozponávania je vidieť na obrázku 9–3.



Obr. 9–3 Závislosť počtu vzorov od ich dĺžky získaných z merania a vyskladávania pomocou Houghovej transformácie [35].

Porovnaním výsledkov získaných vyskladávanou sumačnou metódou 9–2 a výsledkov na obrázku 9–3 je podľa tvaru a sklonu krivky vidieť, že výsledky sú kvalitatívne zhodné použitím kombinácie viacerých algoritmov vyskladávania a viacerých metód rozpoznávania vzoru.

10 Záver

Účelom práce bola analýza simulovaných dát JEM-EUSO experimentu s cieľom určenia pravdepodobnosti nájdania falošnej udalosti. Pod falošnou udalosťou sa rozumie šumom vytvorený signál zameniteľný so signálom tvoreným spŕškami kozmického žiarenia. Simulované JEM-EUSO dáta zodpovedali meraniu svetelného pozadia na nočnej strane Zeme JEM-EUSO detektorom so šumivými charakteristikami s Poissonovým rozdelením. Pravdepodobnosť vzniku falošnej udalosti vo forme falošného vzoru s rovnakými charakteristikami ako má spŕška sekundárneho kozmického žiarenia bola skúmaná z hľadiska pravdepodobnosti výskytu rôzne dlhých falošných vzorov. Analýzou bola určená pravdepodobnosť výskytu falošných udalostí, ktorá klesá s dĺžkou falošných vzorov.

Analýza bola prevedená dvomi metódami rozpoznávania vzorov. Prvou metódou bola metóda zhlukovania vyhľadávajúca spŕšku medzi všetkými kombináciami začiatočného a koncového pixelu, čím sa preverili všetky možné výskyty spŕšok. Druhou bola Houghova metóda. Táto metóda hľadá spŕšky pozdĺž priamok pretínajúcich každý pixel pod množinou uhlov, ktoré sa menia krokovo na základe vopred určeného počtu krokov.

Analýza bola prevedená na inkrementálne sa spresňujúcom popise analyzovaných dát, ktorý smeroval krok po kroku k dátam s plnými charakteristikami JEM-EUSO merania. Obe metódy poskytli kvalitatívne rovnaké výsledky. Ukázali, že v generovanom signále pre dĺžku meraní celého JEM-EUSO detektora rádovo v hodinách je možné nájsť falošné vzory s dĺžkou maximálne 6 až 10 pixelov, kde maximálna dĺžka závisí od definície vzoru.

Keďže Houghova metóda v svojom základnom nastavení hľadá všetky nespojité vzory, bola táto metóda využitá aj na analýzu nespojitého vzoru. Nespojité vzory dosahujú dĺžku až približne 25 pixelov. Signál tvorený spŕškami však produkuje spojité vzory, a tak bola analýza nespojitých vzorov prevedená pre verifikáciu funkčnosti

spúšťacej schémy detektora. Houghova metóda vo verzii bez podmienky spojitosti našla v skúmanom signále JEM-EUSO simulovaných dát charakteristický znak spúšťacej schémy, čím sa potvrdila funkčnosť simulovaných spúšťačov.

Maximálna dĺžka nájdeného falošného spojitého vzoru bola 6 až 10 pixelov podľa definície vzoru. Keďže veľkosť priemetu pixelu na Zemský povrch je 0.51 kilometra v strede citlivej plochy detektora, tak 10 pixelov dlhá sprška zodpovedá sprške dlhej 5.1 kilometra. Analýza bola vykonaná na dátach zodpovedajúcich rádovo hodinám merania celého JEM-EUSO detektora. Pri väčšom množstve dát bude maximálna dĺžka nájdených vzorov rásť podľa zistenej pravdepodobnostnej charakteristiky. Maximálna dĺžka falošných spojitých vzorov pre trojročné meranie tak zodpovedá sprškam so zenitovými uhlami maximálne 30 stupňov.

Charakteristiky reálnych spršok vytvorených UHECR časticami umožňujú určenie zenitového uhla s presnosťou lepšou ako 2,5 stupňa pre častice so zenitovým uhlom väčším ako 30 stupňov. Minimálne 2,5 stupňová presnosť je potrebná pre dosiahnutie základného cieľa misie, ktorým je určenie zdrojov UHECR. Keďže falošný vzor nemôže mať zenitový uhol väčší ako 30 stupňov, nebudú falošné vzory ani čiastočne zasahovať do analýzy reálnych UHECR spršok.

JEM-EUSO kolaboráciou používaná PWISE metóda oddeľujúca signál spršok od šumu pozadia nie je vhodná na analýzu samotného šumového pozadia. PWISE analyzuje len body nad stanoveným prahom, hranicou získanou z analýzy signálu reálnych spršok. Aplikácia tohto prahu na šum pozadia vedie k selekcii minimálneho počtu pixelov, ktoré nikdy nevytvoria vzor dlhší ako 2-3 pixely. Preto PWISE neposkytne pravdepodobnostnú charakteristiku vzniku rôzne dlhých falošných signálov. Z tohto dôvodu boli vyvinuté alternatívne metódy diskriminácie signálu zo šumu pozadia umožňujúce vytvorenie predikcie pre vyššie štatistiky generovaných JEM-EUSO dát.

Použitie dvoch komparatívnych metód rozpoznávania vzorov prinieslo kvalitatívne

podobné výsledky a potvrdilo správnosť výsledkov prevedenej analýzy. Po potvrdení správnosti prevedenej analýzy sa metóda zhlukovania implementovala do kolaboračného rámca ESAF.

Analýza pravdepodobnosti výskytu falošných spíšok poskytla konkluzívne výsledky a ako úloha slovenskej JEM-EUSO skupiny tak dospela k svojmu úspešnému zavŕšeniu.

Literatúra

- [1] BLASCHKE, Filip: Analýza korelovaných spršek kosmického záření: Diplomová práce. Opava: SLU FPF, 2009. 116s.
- [2] ABREU P. et al.: 2011. *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, arXiv:1107.4804 [astro-ph.IM]
- [3] NAGANO, M. WATSON, A.: 2000. *Observations and implications of the ultrahigh-energy cosmic ray* American Physical Society. Jul 2000.
- [4] The Akeno Observatory: Akeno Giant Air Shower Array [online]. Japan: University of Tokyo, 2003. [s.a.] [cit. 2013-11-06]. Dostupné na internete: <<http://www-akeno.icrr.u-tokyo.ac.jp/AGASA/>>.
- [5] Pierre Auger Observatory: The Pierre Auger Cosmic Ray Observatory [online]. Argentina, 2005. [s.a.] [cit. 2013-11-06]. Dostupné na internete: <<http://www.auger.org/index.html/>>.
- [6] Telescope Array: Telescope Array Project [online]. Utah: University of Utah, 2010. [s.a.] [cit. 2013-11-06]. Dostupné na internete: <<http://www.telescopearray.org/>>.
- [7] Luis Anchordoqui et al.: Ultrahigh Energy Cosmic Rays: The state of the art before the Auger Observatory Boston: Department of Physics, Northeastern University, 2002. 5 - 6s.
- [8] ABBASI, R. ABU-ZAYYAD, T. et al.: First Observation of the Greisen-Zatsepin-Kuzmin Suppression. The High Resolution Fly's Eye Collaboration, 2008, arXiv:astro-ph/0703099v2 [astro-ph.IM]
- [9] TORRES, D. BOLDT, E. et al.: Nearby quasar remnants and ultra-high

-
- energy cosmic rays. Phys.Rev. D66, 2002, arXiv:astro-ph/0204419v2 [astro-ph.IM]
- [10] STANEV, Todor: CosmicRays and Exensive Air Showers [online]. Delaware: University of Delaware, 2009. [s.a.] [cit. 2013-11-05]. Dostupné na internete: <<http://indico.cern.ch/event/41547/session/10/material/slides/1?contribId=6/>>.
- [11] Jelley,J.V.: Cerenkov Radiation And Its Applications [online]. Osmania: Osmania University, 1958. [s.a.] [cit. 2013-11-05]. Dostupné na internete: <<https://archive.org/details/cerenkovradiatio030980mbp/>>.
- [12] ARQUEROS,F. BLANCO, F. ROSADO, J.: Analysis of the fluorescence emission from atmospheric nitrogen by electron excitation, and its application to fluorescence telescopes. [online]. Madrid: Universidad Complutense de Madrid, 2009. [s.a.] [cit. 2013-11-05]. Dostupné na internete: <<http://iopscience.iop.org/1367-2630/11/6/065011/>>.
- [13] J.H Adams, et al.: An evaluation of the exposure in nadir observation of the JEM-EUSO mission. Elsevier: JEM-EUSO Collaboration, 2013.
- [14] KAJINO, F. CASOLINO, M. et al.: The JEM-EUSO Instruments 2013, Rio De Janeiro, ICRC, 2013, arXiv:1307.7071v1 [astro-ph.IM]
- [15] 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/>>.
- [16] HACHISU, Y. UEHARA, Y. et al.:Manufacturing of the TA-EUSO and the EUSO-Balloon lenses, Rio De Janeiro, ICRC, 2013, arXiv:1307.7071v1 [astro-ph.IM]
- [17] ADAMS, J. ANCHORDOQUI, L. et al.:Summary Report of JEM-EUSO
-

-
- Workshop at KICP in CHicago, Chicago, 2012, arXiv:1203.3451v1 [astro-ph.IM]
- [18] SANTANGELO, A. PICOZZA, P. et al.: Status of the JEM-EUSO Mission, Rio De Janeiro, 2013, arXiv:1307.7071v1 [astro-ph.IM]
- [19] PASTIČÁK Blahoslav, BOBÍK Pavol, KUDELA Karel: JEM-EUSO experiment for extreme energy cosmic ray observation. In: Physics in Collision, Košice: SAV, 2013 arXiv:1212.6520v1 [astro-ph.IM]
- [20] JAIN A. K., DUIN R. P. W., MAO J.. 2000. *Statistical pattern recognition: a review* In: IEEE Trans. Pattern Analysis Machine Intell., 2000
- [21] Dis Ticaret Ltd.: Pattern Classification [online]. Dis Ticaret Ltd., 2013 [cit. 2013-12-12]. Dostupné na internete: <http://www.byclb.com/TR/Tutorials/neural_networks/ch1_1.htm/>.
- [22] POLIKAR, R. 2006. *Pattern Recognition*. Glassboro, New Jersey : Rowan University, 2006
- [23] Duda, R. Hart, P. Stork, D. 2000. *Pattern Classification*. , New York: Willey, 2000, ISBN: 978-0-471-05669-0
- [24] THEODORIDIS, S., KOUTRUMBAS, K.. 2009. *Pattern Recognition*. Londýn: Elsevier Inc., 2009, 967 s. ISBN: 978-1-59749-272-0
- [25] Wikimedia Foundation, Inc.: Cluster analysis [online]. Wikimedia Foundation, Inc., 2013 [cit. 2013-04-24]. Dostupné na internete: <http://en.wikipedia.org/wiki/Cluster_analysis/>.
- [26] MERNIK, T. GUZMAN, A. et al.: ESAF-Simulation of the EUSO-Balloon In: THE ASTROPARTICLE PHYSICS CONFERENCE in 2013, Rio De Janeiro, The JEM-EUSO Collaboration, 2013
- [27] MERNIK, T. GUZMAN, A. FENU, F. et al.: Simulating the JEM-EUSO
-

-
- Mission: Expected Reconstruction Performance In: THE ASTROPARTICLE PHYSICS CONFERENCE in 2013, Rio De Janeiro, The JEM-EUSO Collaboration, 2013
- [28] Wikimedia Foundation, Inc.: Hough transform [online]. Wikimedia Foundation, Inc., 2013 [cit. 2013-04-25]. Dostupné na internete: <http://en.wikipedia.org/wiki/Hough_transform#Theory/>.
- [29] MERNIK, T. 2009. *Thomas Mernik Thesis* In: Physica Status Solidi (a), Tübingen, 2009, pp. K 87–90
- [30] FENU, F. 2008. *Francesco Fenu Thesis* In: A simulation study of space based missions for Ultra High Energy Cosmic Ray search, Tübingen, 2008
- [31] The ROOT Team: About ROOT [online]. The ROOT Team, 2014 [cit. 2013-11-26]. Dostupné na internete: <<http://root.cern.ch/drupal/content/about/>>.
- [32] EBISUZAKI, Toshikazu et al.: Report on the Phase A Study 2010. Riken: JEM-EUSO Collaboration, 2010. 437 s. JEUSO-110025-01-E-TR-ZZZ
- [33] Berat C. et al.. 2010. *Full Simulation of Space-Based Extensive Air Showers Detectors with ESAF* Italy,France: February 2010, arXiv:0907.5275v3 [astro-ph.IM]
- [34] Bayer Jörg. 2011. *Development of a Cluster Control Board for the JEM-EUSO Mission*: Diplomová práca. Tubingen: Eberhard Karls University, 2011. 27 - 34 s.
- [35] Staroň Miroslav. 2013. *Algoritmy pre rozpoznávanie obrazcov pri štúdiu kozmického žiarenia ultravysokých energií v rámci JEM-EUSO experimentu*: Diplomová práca. Košice: Technická univerzita v Košiciach, 2013.
-

Zoznam príloh

Príloha A Používateľská príručka

Príloha B Systémová príručka

Príloha C DVD nosič

Príloha D Rozšírený abstrakt