

DOI: 10.51981/2588-0039.2023.46.031

СОВРЕМЕННЫЕ МЕТОДЫ АНАЛИЗА ДАННЫХ НА ПРИМЕРЕ ОБНАРУЖЕНИЯ И РЕКОНСТРУКЦИИ ТРЕКОВЫХ СОБЫТИЙ УФ ДЕТЕКТОРА «ВЕРХНЕТУЛОМСКИЙ»

Р. Сараев, С. Шаракин

Научно-исследовательский институт ядерной физики имени Д.В. Скобельцына

Аннотация

В работе представлены методы поиска и реконструкции трековых событий, основанные на машинном обучении. Алгоритм поиска представляет собой нейросетевой триггер, настроенный на выделение характерного пространственно-временного паттерна на сложнопеременном фоне. Байесовская реконструкция позволяет наиболее полно учесть информацию, имеющуюся у исследователя относительно изучаемого явления и самого процесса измерения. Апробация методов осуществлена на данных, зарегистрированных многоканальным детектором, работающем вблизи Мурманска в обсерватории «Верхнетуломская» с сентября 2021 г.

1. Введение

Для изучения быстропеременных (транзиентных) процессов в атмосфере Земли в НИИЯФ МГУ в течение длительного времени разрабатываются и создаются изображающие детекторы, сенсорная поверхность которых составлена из многоанодных фотоэлектронных умножителей (МАФЭУ) [1], [2]. К основным преимуществам таких приборов относится их высокая чувствительность и высокое временное разрешение.

С 2021 года вблизи Мурманска в обсерватории «Верхнетуломская» (68.63 N, 31.78 E) был установлен и начал работать первый из детекторов проекта PAIPS (от англ. Pulsating Aurora Imaging Photometers Stereoscopic System), основной научной задачей которого является изучение пульсирующих полярных сияний, рис. 1. Линзовый объектив детектора строит изображение УФ-источника (300-400 нм) на многоканальном фотоприемнике, составленном из четырех МАФЭУ (общее число каналов 256, поле зрения одного канала составляет около 1°). В мониторинговом режиме прибор регистрирует данные с временным разрешением 1 мс (в сезоне 2021-2022 – 40 мс).

Детектор также служит тестовой площадкой для разработки и апробации современных методов анализа данных: поиска событий с характерным пространственно-временным паттерном сигнала и их реконструкции. В настоящей работе приведен пример анализа так называемых трековых событий, т.е. событий, в которых пик сигнала в изображении на фотоприемнике перемещается из канала в канал вдоль определённого направления. В качестве трековых выступают хорошо идентифицируемые события типа движения звёзды по небосклону, пролет через поле зрения прибора спутника или самолета, метеоры, а также релятивистская «экзотика»: космические лучи, пучки энергичных электронов, нуклеариты и др.

Типичный цикл работы с данными изображающего детектора можно разбить на несколько этапов: 1) пред-просмотр и пред-обработка полученных данных; 2) поиск и распознавание событий; 3) реконструкция модельных параметров события; 4) интерпретация полученных результатов.

На этапе предобработки требуется сконвертировать измерения в данные с наиболее подходящим для анализа интервалом дискретизации, учесть неоднородность распределения чувствительностей по каналам фотоприемника, при необходимости применить временные фильтры к сигналу в каждом канале и/или пространственные к изображению.

На этапе поиска основные трудности связаны с непростой структурированностью сигнала на сложном быстроменяющемся фоне. Для решения этой задачи было принято решение использовать нейросетевой триггер.

На этапе реконструкции важно иметь возможность постепенного усложнения используемой модели с тем, чтобы она могла включать в себя всё больше особенностей как самого явления, так и процесса измерения. Наиболее подходящим для этого, с нашей точки зрения, является реконструкция в рамках вероятностного (байесова) вывода [3].

Для удобства анализа зарегистрированных данных нами было создано специальное приложение с графическим интерфейсом, включающим инструменты для пред-просмотра событий, конвертации его в разные временные разрешения, выравнивания чувствительностей. Специальные инструменты приложения позволяют осуществлять ручную и полу-автоматическую разметку для формирования обучающей выборки

триггера, конструировать и модифицировать архитектуру используемой нейросети, а также непосредственно осуществлять с их помощью поиск трековых событий.

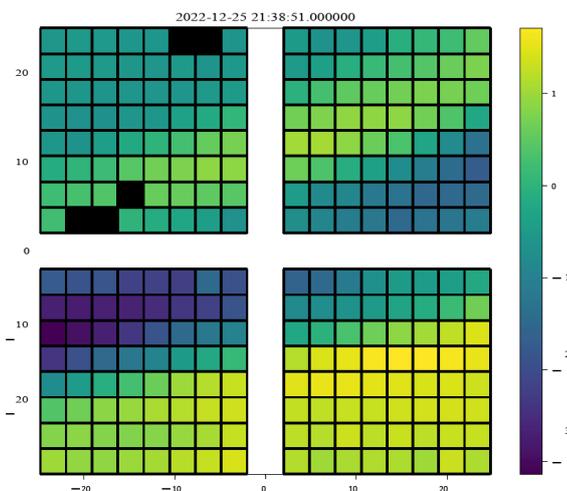


Рисунок 1. Пример полярного сияния в данных детектора «Верхнетуломский».

2. Поиск трековых событий

Фоновый сигнал, на котором требуется идентифицировать трековое событие, является сильно неоднородным в пространстве и переменным во времени. Кроме того, время от времени значительная часть поля зрения прибора заполняется кратковременными, но яркими вспышками. В связи с этим на этапе предобработки осуществляется трехэтапное преобразование сигнала: 1) выделение активного сигнала из фона с помощью фильтра скользящей медианы; 2) выравнивание флуктуаций путем масштабирования на скользящую медиану абсолютного отклонения (MAD); 3) вычитание мгновенного медианного значения по всем каналам фотоприемника. (Медианные значения применяются для придания робастности процедуре.)

Для поиска трековых событий используется бинарный классификатор (триггер), представляющий собой нейронную сеть. Ее алгоритм рассматривает не интегральное изображение трека (как, например, в работах [4]), а его полный пространственно-временной паттерн, представленный совокупностью временных развёрток активных сигналов. Для повышения чувствительности триггер настраивается на определенный диапазон скоростей, с которыми пик сигнала перемещается по каналам фотоприемника. Это подразумевает предварительную конвертацию данных в соответствующее временное разрешение.

Архитектура сети включает в себя два свёрточных слоя, один из которых осуществляет пространственную свёртку, другой – временную. Аппроксимация трехмерного слоя (32x5x5) двумя слоями увеличивает число нелинейностей в 2 раза, что ускоряет обучение сети [5]. Выход свёрточных слоев передается на слой с выбором максимума (max pooling) и три полносвязных слоя. Алгоритм реализован на языке python на библиотеках tensorflow+keras.

Трековые события довольно просто смоделировать, используя низкоразмерные параметрические модели. С другой стороны, сложно-структурированный и многообразный фон слишком сложен для моделирования, помимо обычного шума он может содержать облака, полярные сияния, вспышки и т.п. Поэтому элементы обучающей выборки синтезируются из реального фона и модельных треков. С помощью инструмента ручной разметки были выделены интервалы фонового сигнала из четырёх ночей, содержащих облака, полярные сияния и ясное небо. Модельные треки выбираются из предварительно созданной базы данных или генерируются на лету с разными значениями ключевых параметров (скоростью, направлением, яркостью, формой кривой свечения).

Инструмент формирования синтезированной выборки и обучения нейросети показан на рисунке 2 (настройка компонентов нейросети происходит по кнопке Reset model).

Тестовая выборка объемом 12 000 событий создавалась аналогичным образом из данных, не участвовавших в обучении. В процессе обучения нейросеть выходит на accuracy (доля правильно определённых явлений) 97%: триггер верно идентифицировал 5446 треков (из 5992) и 5801 сложных фонов (из 6008). Большинство ложноотрицательных событий лежит на границе интервалов по амплитуде сигнала и по скорости в обучающей выборке. В ложноположительных событиях участок полярного сияния имитирует трек.

Таблица 1. Результаты реконструкции спутникового события CZ-6A DEB

Параметр	$\Phi_0, ^\circ$	U_0 , пикс/такт	E_0	σ_{psf} , пикс
МАФЭУ А	-96.6	0.075	32.8	0.15
МАФЭУ С	-99.7	0.082	59.1	0.20

4. Вывод

В работе для поиска и реконструкции трековых событий использованы методы машинного обучения. Алгоритм поиска, основанный на сверточной нейросети, может быть расширен для выделения других явлений, таких как полярные сияния и микровсплески, вызванные высыпанием релятивистских электронов. Байесовский подход, использованный в реконструкции, позволяет обобщить метод для восстановления трёхмерной кинематики метеоров.

Исследование выполнено при финансовой поддержке Российского научного фонда, грант № 22-62-00010 (<https://rscf.ru/project/22-62-00010>).

Литература

1. P.A. Klimov et al. "The TUS detector of extreme energy cosmic rays on board the Lomonosov satellite". Space Sci. Rev., 2017, V. 212, Is. 3-4, P. 1687-1703. <https://doi.org/10.1007/s11214-017-0403-3>
2. S. Bacholle et al. "Mini-EUSO Mission to Study Earth UV Emissions on board the ISS". The Astrophysical Journal Supplement Series, 2021, V. 253, Is. 2, Art. 36. <https://doi.org/10.3847/1538-4365/abd93d>
3. D.S. Sivia and J. Skilling. Data Analysis - A Bayesian Tutorial. 2nd. Oxford Science Publications. Oxford University Press, 2006.
4. Miles Winter et al. "Particle identification in camera image sensors using computer vision". Astroparticle Physics, 2019, V. 104, P. 42-53. <https://doi.org/10.1016/j.astropartphys.2018.08.009>
5. D. Tran et al. "A Closer Look at Spatiotemporal Convolutions for Action Recognition". IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018, P. 6450-6459. DOI: 10.1109/CVPR.2018.00675
6. J. Salvatier, T.V. Wiecki and C. Fonnesbeck. "Probabilistic programming in Python using PyMC3". PeerJ Comput. Sci., 2016, 2, e55. <https://doi.org/10.7717/peerj-cs.55>