



107
сообщения
объединенного
института
ядерных
исследований
дубна

6507/83

P10-83-612

Б.Нергуй, Г.А.Ососков

ОСНОВНЫЕ АЛГОРИТМЫ
И СТРУКТУРА ПРОГРАММЫ **FILTR**

1983

Введение

Для обработки снимков α -р взаимодействий, экспонированных на 1-м жидководородной пузырьковой камере ЛВЭ ОИЯИ, при импульсе пучка свыше 10 ГэВ/с с использованием спирального измерителя (СИ) ОИЯИ/1/, потребовалось решить комплекс задач, среди которых – разработать систему программ для офф-лайн обработки данных спирального сканирования указанных снимков и, прежде всего, программы фильтрации, осуществляющие начальную обработку данных.

Необходимость в процессе фильтрации результатов машинного сканирования является общей чертой автоматических измерительных систем, которые вместе с функцией измерения вынуждены были перенять у человека также и функции распознавания измерительного объекта и организации процесса измерений.

С этой целью программы фильтрации практически любого сканирующего автомата должны выполнять следующие операции:

1. Контроль входных данных.
2. Выделение точек, относящихся к проекции данного события, и отсев посторонних точек, порожденных случайными шумами, царапинами на планке, а также треками, не входящими в событие.
3. Сжатие выделенных, полезных данных путем усреднения для получения 10+20 мастер-точек на каждый трек.
4. Сборка, окончательная проверка данных о всех проекциях текущего события и формирование записи о нем на магнитной ленте в согласованном формате для последующей обработки по геометрической и кинематической программам.

Выбор прибора, сканирующего изображение в полярной системе координат радиально ориентированной щелью по спирали от вершины, измеряемой оператором в декартовой системе координат, почти однозначно определил набор алгоритмов для выполнения второго, главного из перечисленных выше пунктов фильтрации. Этот набор, ставший традиционным для устройств данного типа /2-4/, состоит из алгоритмов, выполняющих три последовательных этапа обнаружения треков. Вначале для поиска возможных кандидатов в треки, исходящих из вершины, применяется метод поворотных гистограмм /5/, хорошо работающий в районе первых 10+15 витков спирали, где существенным образом сказывается влияние фильтрующего эффекта щели и характерного растяжения координатной плоскости при прямоугольном представлении полярных координат (см. рис.1).

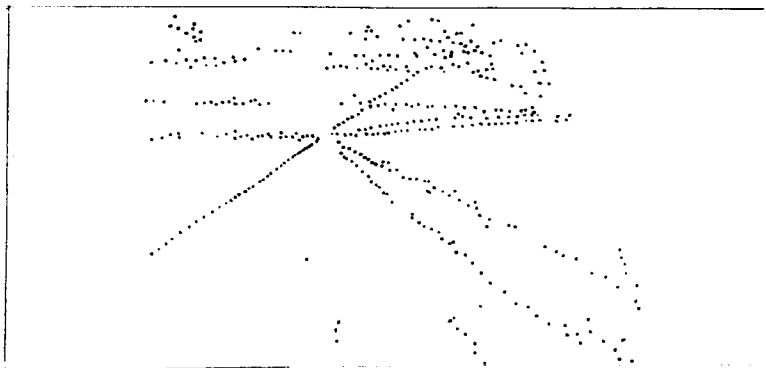


Рис. 1а

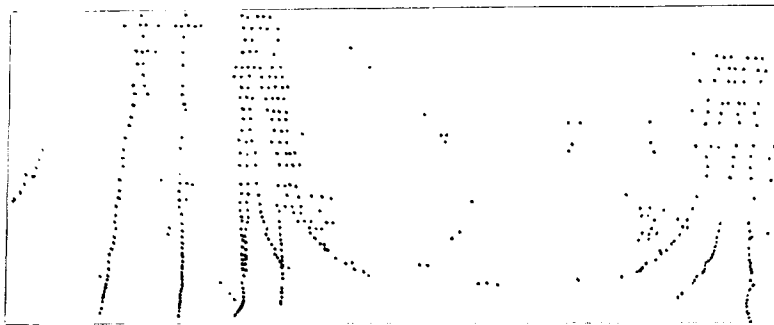


Рис. 1б

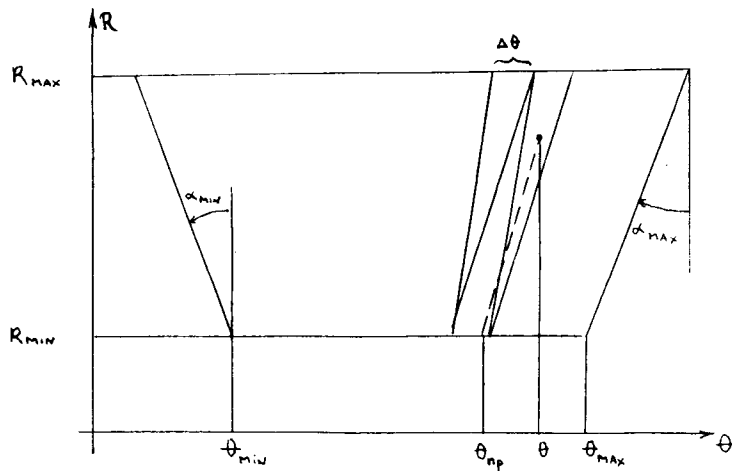


Рис. 2

Затем производится прослеживание полученных кандидатов в треки путем многократного применения подгонки модельного уравнения трека по методу наименьших квадратов (м.н.к.) как для предсказания положения трека на следующем витке спирали, так и для проверки принадлежности треку точки, найденной в области предсказания. В заключение отслезенные треки сопоставляют с так называемыми опорными точками (ОПТ), которые в необходимых случаях измеряются оператором в декартовой системе координат.

Эта схема оказалась достаточно надежной в применении, но громоздкой с точки зрения затрат ресурсов ЭВМ. Только на этапе гистограммирования тратилось около 40% всего времени фильтрации ^{/2/}, причем на простейшую программную реализацию метода поворотных гистограмм требовалось от 5 до 8 тыс. ячеек памяти ЭВМ. Поэтому желание упростить эту традиционную схему, уменьшить ее громоздкость с вычислительной точки зрения, и в то же время максимально использовать априорную информацию, привело к появлению второго подхода к решению задачи обнаружения треков, состоящего в объединении первого и второго этапов с помощью метода шнуров (stringing) ^{/5/}, причем начинать предлагалось сразу с третьего этапа, т.е. вначале прослеживать от наличных ОПТ, а уже потом, при необходимости, от вершины. Однако исследование ^{/6/} показало недостаточную надежность метода шнуров как обычного, так и усовершенствованного, при прослеживании треков, пересекающихся под малыми углами. Этот недостаток удалось преодолеть на основе байесовского подхода, позволившего полнее учитывать предысторию трека по уже отслезенным точкам ^{/7/}.

Тем не менее авторы настоящей работы остановились на традиционной схеме выделения событий, т.к. прежде всего удалось создать быстрый алгоритм метода поворотных гистограмм, втрое снижающий затраты времени и в 10 раз - память ЭВМ ^{/8/}. Наряду с другими мерами по повышению эффективности процесса фильтрации (см. ^{/9/}), это позволило значительно сократить разрыв в трудоемкости программ, основанных на традиционном и байесовском методах. Было учтено также, что последний был опробован лишь на d-p событиях сравнительно простой топологии ^{/7/}, в то время как рассмотрение специфики данных α -p взаимодействий при энергиях свыше 10 ГэВ показывало на значительное усложнение условий распознавания, вызванное особым характером пучковых треков, наличием "узких" пар сливающихся вторичных треков, близких по кривизне и общей для взаимодействий при высоких энергиях проблемой плохой различимости треков вторичных частиц, выходящих вперед в узком конусе.

При создании программы фильтрации было также необходимо учесть отличия СИ ОИЯИ от аналогичных западных приборов: меньшую разрешающую способность по углу и радиусу (15 разрядов для R и 16 - для θ),

отсутствие цифровой информации об амплитуде и ширине импульса сигналов отсчетной системы и, с другой стороны, — наличие признака e_{full} — пересечения щелью опорной линии $/I/$, позволяющего однозначно разбить данные сканирования по виткам спирали. Помимо влияния на программу обнаружения, эта специфика СИ ОИЯИ определила этап контроля входных данных, производимый в процессе их перепаковки из формата управляющей ЭВМ во внутренний формат FILTR'a.

Кроме этого, оригинальность отсчетной системы СИ ОИЯИ $/I/$ потребовала разработки специальной калибровочной программы $/IO, II/$, вычисляющей калибровочные параметры, необходимые для взаимных переходов между полярной и декартовой системами координат.

Ограничения, накладываемые тем обстоятельством, что математическое обеспечение СИ ОИЯИ должно было стать типовым, пригодным для институтов, объединившихся в коллаборацию пользователей СИ и снабженных только ЭВМ БЭСМ-6, практически лишило возможности использования преимуществ таких ЭВМ, как CDC-6600 или ICL-1906A.

В силу всех этих обстоятельств авторы не могли использовать программы фильтрации других систем, аналогичных СИ, в том числе и первый вариант программы FILTR $/I2/$, несмотря на его усовершенствования $/I3/$.

При сохранении традиционной трехэтапной схемы выделения треков события были практически заново написаны основные алгоритмы и перестроена структура как программы, так и способа организации данных, что позволило поднять эффективность программы FILTR и сделать ее пригодной для обработки снимков α -р взаимодействий, экспонированных на μ -м водородной камере ОИЯИ при энергиях пучка 8,5 ГэВ и выше.

Программа была передана в с е к т о р С И 30 апреля 1982 г. после автономной отладки и набора некоторой статистики.

Ниже приводится краткое изложение основных алгоритмов по этапам фильтрации в соответствии с общими принципами стратегии массовой обработки потоков экспериментальных данных $/I4/$, даются описания оптимальной структуры данных, архитектуры программы FILTR и ее основные характеристики. Указаны также некоторые идеи, которые по разным причинам не были включены в первоначальный вариант программы, но могут оказаться полезными разработчикам аналогичного матобеспечения.

2. Основные алгоритмы

2.1. Контроль данных

Недооценка контроля вводимой информации приводит не только к ненужным затратам на обработку дискретных данных, но, и главное,

к возможному искажению последующих физических выводов. Поэтому в программу FILTR включены проверки соответствия паспортных данных реально вводимой информации, а также наиболее простая и эффективная проверка качества вводимых данных спирального сканирования — контроль монотонности роста радиуса. Это делается в двух местах: сразу же после ввода, проверяющего монотонность роста радиусов R_{full} точек $(R_{full}, \theta_{full})$, содержащих признак пересечения щелью опорной линии $/I/$. При этом для обнаружения возможных кольцевых пропусков данных проверяется также и монотонность роста шагов спирали.

Вторая проверка монотонности делается уже для радиусов всех точек, принадлежащих области поиска трековых элементов. На этом же этапе, предшествующем обнаружению треков, предполагается ввести также контроль качества измерений реперных крестов, уже опробованной в "байесовом" варианте программы фильтрации $/7/$.

2.2. Обнаружение треков-элементов

Быстрый алгоритм метода поворотных гистограмм $/8/$ позволяет с минимальными затратами машинного времени получить вклады сразу всех точек выбранной области в набор гистограмм, каждая из которых образует характерный "веер" частично перекрывающихся полосок с наклонами, изменяющимися в заданном диапазоне (рис. 2) в зависимости от наименьшего возможного радиуса кривизны треков в данном эксперименте.

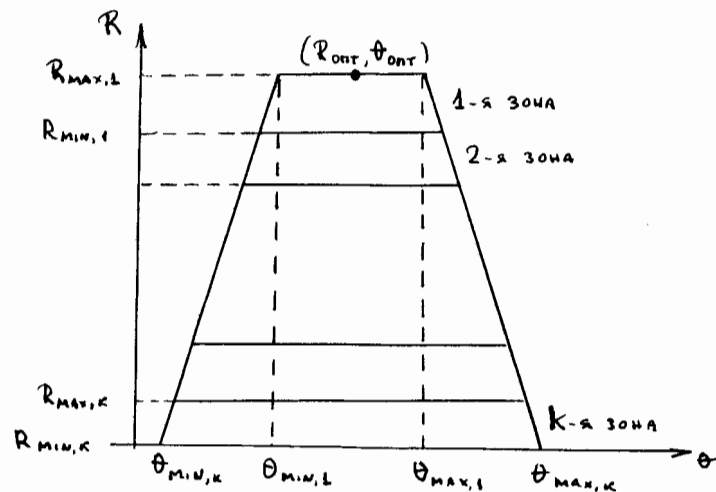


Рис. 3

При превышении содержимого какой-то из ячеек гистограммы порога P из точек, попавших в эту ячейку, формируется банк трек-элемента. Совмещение процедур гистограммирования и формирования банков трек-элементов, осуществленное в алгоритме /8/, позволило добиться высокой точности и экономичности: так, при "веере" из 21 направления и разбиении отрезка оси $\theta(-\pi, \pi)$ на 360 ячеек объем рабочей памяти, потребный для реализации алгоритма, сократился до 380 слов.

Появление дубликатов трек-кандидатов, вызываемое перекрытием соседних полосок гистограмм (это необходимо, чтобы избежать "разрезания" трека вдоль границы между полосками) довольно просто устраняется затем специальной программой.

Для гибкой организации поиска треков на любом участке области сканирования в качестве входных параметров подпрограммы гистограммирования, определяющих зону обнаружения, заданы номера витков спирали, соответствующие R_{\max} и R_{\min} , номера ячеек гистограммы, соответствующие θ_{\max} и θ_{\min} (рис. 2), а также номера первого и последнего наклонов полос гистограмм.

Такая организация процедуры гистограммирования позволяет легко осуществить идею двухэтапного гистограммирования /14/ в тех случаях, когда происходит значительное превышение порога P , означающее, что в одну полоску гистограммы попадают точки нескольких треков. Эти направления должны гистограммироваться повторно с вдвое меньшим шагом по углу θ).

При задании зоны гистограммирования для первичного поиска треков в районе вершины размер зоны по радиусу выбирается в 10 витков, причем первые 4+5 витков пропускаются, т.к. ближе к вершине возможно полное слияние группы вторичных треков, а также изгибы концов треков у вершины в виде характерных крючков в случаях неточного измерения вершины оператором.

Метод поворотных гистограмм применяется также и для поиска треков в районе тех опорных точек (ОПТ), для которых не хранилось соответствующих треков среди уже найденных. Направление поиска к вершине определяет зону гистограммирования, длина которой при обнаружении трека увеличивается путем окаймления (рис. 3).

Специальная зона гистограммирования используется в случаях, когда при первичном гистограммировании не обнаруживается пучковый трек. Поскольку его направления и кривизна известны заранее, эта зона имеет

вид узкого длинного сектора, идущего вдоль ожидаемого направления пучкового трека.

Во всех случаях, когда гистограммирование ведется по зонам переменной длины вдали от вершины, где шаг спирали также меняется, используется динамическое задание порога P в зависимости от числа оборотов спирали между $R_{\max,1}$ и $R_{\min,k}$ ($\min P=3$).

2.3. Прослеживание треков

Полярное уравнение трека было выведено /15/ как приближенное уравнение окружности, проходящей на малом расстоянии β от полюса

$$\theta(R) = \theta_0 + \alpha R + \beta/R + \gamma R^3.$$

Коэффициенты α и γ в случае окружности связаны соотношением $\gamma = \frac{\alpha^2}{6}$ /3/. Как оказалось, отказ от этой связи дает широкий класс кривых, хорошо описывающий даже треки в больших пузырьковых камерах /4/.

Фактически в зависимости от расстояния R до вершины в программе FILTR предусмотрена динамическая смена математических моделей трека: вблизи вершины $\gamma=0, \beta \neq 0$ и при превышении величиной $|\beta|$ заданного порога трек отвергается; при прослеживании от ОПТ к вершине $\gamma=0, \beta=0$, но с приближением к вершине β "включается". Включение коэффициента γ при прослеживании в области, достаточно далекой от вершины, может быть корректно произведено только после детального исследования характера зависимости между γ и кривизной (2α) вторичных треков для данного эксперимента. Поэтому в настоящем варианте FILTR'a хотя и разработан полностью необходимый программный аппарат для прослеживания с помощью модели (3) со всеми четырьмя параметрами, но для его использования требуется добавить проверку соответствующего условия. Пока же параметр γ применяется лишь при финальной подгонке трека.

Прослеживание начинается с того, что к каждому из трек-элементов, оставшихся после устранения дубликатов, подгоняется соответствующая модельная кривая путем минимизации функционала

$$S^2 = \sum_{i=1}^k w_i [\theta_i - \theta(R_i)]^2. \quad (4)$$

Веса w_i выбраны пропорциональными R_i^2 , чтобы обеспечить декартову (а не угловую) метрику при измерении расстояния точки от модельного уравнения.

Полученная кривая используется для предсказания положения θ_j^x -точки трека на следующем витке спирали (при этом существенно

*) Метод пока не включен в программу FILTR.

используются точные значения шага спирали в данной области предсказания). При обнаружении в некоторой окрестности прогноза e_j^* точки (наиближайшей к ней, если их окажется несколько), эту найденную точку включают в банк точек трека и перевычисляют параметры, чтобы проверить, не превысит ли среднеквадратичный разброс $\sigma = \sqrt{s_{\min}^2 / (n-2)}$ заданную константу, определяемую максимально возможной шириной трека. В случае превышения происходит выбор далеко отстоящих точек.

При минимизации функционала (4) коэффициенты нормальных уравнений не перевычисляются заново при добавлении и удалении точек, а корректируются с учетом координат этих точек. Поиск в окрестности прогнозируемой точки осуществляется по двум дорожкам: более узкая соответствует ширине трека на пленке, а вторая - удвоенной ширине. При обнаружении точек в узкой дорожке точки, найденные в широкой дорожке, не рассматриваются. В случае же ненахождения точки в узкой дорожке увеличивается специальная функция пропуска $G^{1/2}$, зависящая от наличия или отсутствия точек трека на трех последних витках спирали.

Наличие точки только в широкой дорожке может быть вызвано либо неточным прогнозированием из-за пропусков на нескольких предыдущих витках спирали, либо уходом аппроксимирующей кривой из-за присоединения к треку ложных отсчетов. Поэтому после анализа пропуска процесс прослеживания продолжается с присоединением к треку точки в широкой дорожке с удалением (если пелел этим не было пропуска в узкой дорожке) последней из присоединенных точек.

2.4. Обработка данных об опорных точках

После прослеживания всех трек-элементов устанавливается их соответствие с опорными точками (ОПТ), если они были измерены оператором. С этой целью декартовы координаты ОПТ с помощью данных калибровки переводятся в полярные координаты и вычисляются взаимные дуговые расстояния между всеми ОПТ и треками. Для каждой из ОПТ находится ближайший к ней трек и ОПТ "приписывается" к нему при выполнении двух условий:

- расстояние трека до ОПТ не превышает утроенной ширины трека;
- расстояние DR_j по радиусу от последней точки трека до ОПТ меньше заданной константы.

Если DR_j оказывается отрицательной, то производят дополнительный поиск от конца трека до ОПТ и далее.

Если ОПТ не приписется ни к одному треку, производится гистограммирование в направлении к вершине по окаймляющим зонам, что,

как правило, позволяет найти трек, даже если он отдален от ОПТ разрывом, вызванным бликом или областью слабого контраста. Для случаев неудачного поиска треков от ОПТ (возникающих, например, когда повторно находится один из уже найденных треков в узкой паре) в ^{14/} была предложена процедура NEWPOL, состоящая в перенесении полюса в ОПТ, что обеспечивало резкое увеличение угловых расстояний в этом районе и, соответственно, лучшую селективность метода поворотных гистограмм. При проверке на снимках с α -P взаимодействий метод NEWPOL показал высокую надежность и поэтому будет включен в программу FILTR.

При спиральном сканировании возможен ряд обстоятельств (полное наложение следов, разрыв близко идущих треков из-за блика на пленке, слишком короткий или слишком кривой трек), когда даже постановка одной ОПТ не спасает положения. Поскольку размер массива в памяти управляющей ЭВМ для размещения точек, измеренных в декартовых координатах, рассчитан максимально на 18 точек, то естественна мысль о применении неиспользованных точек для проведения в этих обстоятельствах ручных домеров, т.е. измерения на таких треках оператором нескольких специальных ОПТ (с признаком "0"). Чтобы было можно без специальных геометрических сопоставлений отличать домеры, относящиеся к различным трекам, приняты следующие ограничения: на треке, измеренном вручную, должно быть поставлено не менее трех ОПТ, причем последняя от вершины должна иметь либо признак "6" обычной ОПТ, либо отмечать остановку (признака "2") или излом ("3") трека, либо быть вторичной вершиной ().

По желанию пользователя точки, измеренные вручную, можно использовать двойко: либо применять их как маску для последующего поиска данных сканирования в дорожке вокруг этой маски, либо прямо включить эти точки в число мастер-точек трека. Последняя возможность позволяет в принципе обрабатывать события, измеренные целиком в ручном режиме, но для этого потребуются небольшие переделки в программе FILTR, обусловленные соответствующими изменениями форматов входных данных и необходимыми добавлениями в блоке управления фильтрацией.

2.5. Идентификация треков-кандидатов и сборка событий

В процессе фильтрации могут возникать следующие возможные ситуации, в которых приходится решать проблему выбора одного из нескольких треков-кандидатов, близких друг к другу:

- а) один и тот же трек найден дважды;
- б) вблизи ОПТ найдены несколько треков;
- в) несколько треков найдены в области пучковых треков.

Сложность решения вызвана наличием нескольких, чаще всего противоречивых требований, которым одновременно должен удовлетворять "хороший" трек. После исследования на большой статистике были отобраны следующие критерии, определяющие качество трека: 1) максимум N - числа точек трека; 2) минимум β - расстояния от вершины или R_1 - радиуса первой точки трека; 3) минимум σ - среднеквадратичного разброса точек трека; 4) минимум D - дугового расстояния от трека до ОПТ.

Для пучкового трека (задача В), среднее значение кривизны $\bar{\alpha}$ которого известно заранее, предлагается еще один критерий: 5) минимум $|\alpha - \bar{\alpha}|$. Учет физики определенного эксперимента дает дополнительные важные критерии: например, в α -р эксперименте возникает требование (6) к количеству вторичных треков I_{TOP} : если $I_{TOP} \leq 3$, то все $\alpha < 0$, если $I_{TOP} \geq 4$, то по крайней мере для одного из вторичных треков $\alpha > 0$. В задаче В еще до проверки критериев I+5 выбрасывают треки, не удовлетворяющие условию $\alpha_{LOW} < \alpha < \alpha_{HIGH}$ (α_{LOW} , α_{HIGH} - границы заданного интервала вокруг $\bar{\alpha}$), что дает иногда заметное сокращение треков кандидатов. Проверка критериев I+3 и 5 делается с помощью минимизации на множестве треков-кандидатов функционала

$$F = - (w_1 \cdot N + \frac{w_L}{|\beta| R_1} + \frac{w_3}{\sigma} + \frac{w_4}{|\alpha - \bar{\alpha}| + 10^{-4}})$$

где веса w_i подобраны по обучающей выборке вторичных треков α -р эксперимента при 8,5 ГэВ. Для задачи А $w_4 = 0$.

В задаче В исследование показало, что число точек трека не является столь существенным фактором, как близость трека к ОПТ (расстояние D) и то, как далеко он был прослежен к вершине (R_1). Поэтому зависимость функционала от R_1 была усилена

$$F = - (\frac{wS_1}{D} + \frac{wS_2}{|\beta|} + \frac{wS_3}{\sigma} + \frac{wS_4}{R_1^2})$$

Веса wS_i в задаче В были подобраны такими:

$$wS_1 = 1.0, wS_2 = wS_4 = 0.1; wS_3 = 0.5.$$

Данные оцифровки, вошедшие в каждый из треков, подлежат затем усреднению для получения не более 16 мастер-точек, которые переводятся затем с помощью калибровочных преобразований в декартову систему координат и компонуется по проекциям события в соответствии с входным форматом программы GEOKIN (см., например, /16/).

Усреднение для получения мастер-точек производится по формулам из работы /3/. В программе FILTR предусмотрен контроль выходной информации, не пропускающий к дальнейшей сборке проекции, если в них отсутствует пучковый трек.

Следует заметить, что при самой тщательной проверке качества треков эта проверка проводится в плоскости проекции и не может гарантировать от принципиальной возможности появления числа вторичных треков, большего, чем предусмотрено топологией. Более того, ужесточение требований к трекам может привести к отсеву истинных проекций треков, имеющих, например, больший среднеквадратичный разброс. Корректный выбор проекций треков из нескольких кандидатов может быть сделан только с помощью специальной процедуры трек-матч с выходом в пространство, на наличие которой в программах последующей обработки и рассчитан выход программы FILTR.

Параметр, определяющий возможное число вторичных вершин в настоящем варианте FILTR'a, равен 2. При объединении проекций в событие проверяется, чтобы при браковке проекций их не осталось меньше двух, причем разной четности, чтобы избежать случаев параллельности оптической базе. Все случаи нарушения требований контроля сборки сопровождаются диагностикой.

В программе имеется также блок подсчета статистической информации о числе событий, прошедших и не прошедших обработку, об общем числе ОПТ, числе треков, измеренных вручную, и общем числе точек на них и т.д.

5. Структура данных

В программах, предназначенных для сложной обработки больших объемов экспериментальной информации, весьма важными являются вопросы экономии памяти ЭВМ и такой рациональной организации данных, которая обеспечивала бы минимизацию операций над ними для получения достаточно высокой скорости обработки.

Помимо калибровочных и титульных констант, общих для группы событий, информация о каждой проекции, вводимая в FILTR, состоит из координат точек, измеренных в $x-y$ системе, и, главным образом, данных спирального сканирования, упакованных в массиве IDIC по одному слову на точку. Рабочая память программы FILTR, необходимая для выполнения этапов обнаружения и прослеживания треков, состоит из ряда массивов, степень заполнения которых (так же, как и для массива IDIC) может сильно изменяться от события к событию. Сюда входят рабочие массивы для метода поворотных гистограмм, для хранения кандидатов в трек и массив tm для накопления результатов фильтрации. Кроме того, значительное место рабочей памяти необходимо для накоп-

ления различного рода статистической информации, выдачи рисунков на широкую печать и т.д.

Можно выделить два совершенно различных подхода к решению задачи об организации данных. Первый — наиболее простой с программистской точки зрения (см., например, /12/), — заранее извлечь из IDIG и разместить в отдельном массиве пары координат (R, θ) всех точек данной проекции, заполняя этими парами и остальные рабочие массивы по мере продвижения процедур фильтрации. С учетом уже упомянутых неоправданно больших затрат на гистограммирование (до 8 тыс. ячеек) это решение оказалось неудачным, так как с ростом множественности событий и усложнением программы быстро привело к переполнению памяти ЭВМ. Сегментация программы на оверлеи трех уровней /17/ оказалась неэффективной при дальнейшем развитии программы.

Второй подход, предложенный авторами CERN-POON /2/, — упаковка данных. Вся информация спирального сканирования в запакованном виде хранится в массиве IDIG и извлекается по мере необходимости с помощью набора малых автокодных подпрограмм-функций. Это дало возможность проводить фильтрацию, используя только адреса точек в массиве IDIG, упакованные по несколько штук в слове, без существенных потерь счетного времени, т.к. функции распаковки — упаковки полностью размещались в сверхбыстрой стековой памяти CDC-6500.

Промежуточным решением проблемы экономии памяти было использование в апрограме POON — ИФВЭ /4/ системы динамического распределения памяти, подобной системе HYDRA /18/, адаптированной для ЭВМ ICL-1906A.

Мы не могли воспользоваться преимуществами больших ЭВМ, подобных ICL-1906A или CDC-6600. Был проведен тщательный анализ потоков данных и частоты использования каждой из координат в процессе фильтрации, показавший, что проверка связности точек ведется по угловым значениям θ . Они главным образом используются в вычислениях, проводимых в процессе фильтрации, в отличие от значений радиуса, используемых много реже, в основном при финальной подгонке, для которой радиусы извлекаются из массива IDIG в соответствии с индексом θ .

Благодаря возможности предварительно разбить все данные в массиве IDIG на группы по виткам спирали с указанием прироста радиуса на каждом витке, можно ограничить поиск при прослеживании одним витком, устанавливая при необходимости соответствие между номером витка и радиусом с помощью быстрого двоичного поиска.

Подсчет временных затрат показал, что хранить θ (16 разрядов) в упакованном виде для ЭВМ типа БЭСМ-6 или CDC-6500 — нецелесообразно. Таким образом, в программе FILTR в распакованном виде хра-

нятся только координаты θ всех точек проекции, а обе координаты используются лишь для хранения трек-элементов в рабочем массиве RTH на 100 точек, расположенных в порядке роста R , чтобы избежать процедуры перераспределения точек при смене направления прослеживания.

В массив отфильтрованных треков T_M в ходе фильтрации добавляется информация о найденных или вычеркиваются сведения о некачественных треках, так что длина этого массива меняется от случая к случаю. Известна и длина той части памяти, в которой накапливаются гистограммы, но в зависимости от требований пользователя эта часть памяти иногда частично, а иногда полностью не используется в течение данного сеанса обработки. Такая изменяющаяся степень заполнения массивов, характерная для программы фильтрации, приводит к идее об объединении резервных частей этих массивов для более рационального их использования, т.е. фактически к известной идее динамического распределения памяти. Однако использование готовой системы динамического распределения памяти HYDRA для сравнительно небольшого числа рабочих массивов программы FILTR оказалось нерациональным. Это было осуществлено с помощью значительно более простого программного аппарата из 6-8 подпрограмм, динамически распределяющих рабочие массивы без промежутков в одномерном массиве Q , заданном в качестве непомеченного блока COMMON. Во всех подпрограммах, которые используют рабочие массивы, обращение к их элементам происходит по имени массива, т.к. с помощью декларации EQUIVALENCE для них выделена общая память, начиная со второй ячейки Q , в то время как в первой ячейке хранится значение максимальной длины Q .

Все декларации COMMON, DIMENSION, EQUIVALENCE, определяющие рабочую память программы FILTR, записаны в виде макроопределений языка директив UPDATE /20/, выполняющего функции хорошо известной системы программ PATCHY /19/, и с помощью макровывоза UPDATE эта декларированная часть вставляется в соответствующие подпрограммы.

Для достижения наилучшей мобильности программы внутренний формат массива IDIG выбран таким же, как для ЭВМ БЭСМ-6 (48 разрядов).

4. Архитектура и основные характеристики программы FILTR

Основные блоки программы FILTR полностью соответствуют этапам фильтрации и реализуют последовательность алгоритмов, изложенную в п. 2. Управление фильтрацией выполняет блок управления, реализованный в виде подпрограммы LEADER (см. блок-схему на рис. 4), которая иницирует работу отдельных блоков в зависимости от конкретной ситуации, сложившейся в результате работы предшествующего блока.

В архитектуру программы FILTR заложен системный подход, основанный на принципах модульности, иерархичности и максимальной восходящей утилизации модулей (последнее означает стремление к такой организации вызова модулей, т.е. набора входных и выходных параметров, которая давала бы возможность максимально частого их использования, и не только на одном, но и на более высоких уровнях иерархии).

Анализируя процесс обработки данных при фильтрации, можно выделить следующие три уровня обработки:

нижний уровень, где осуществляются манипуляции с типами данных, сложность которых нарастает в ходе фильтрации: упакованные данные оцифровки, координаты, трек-элементы, мастер-точки, заголовки треков, треки, строки рисунков, символы протоколов и диагностических сообщений, элементы массива Q и т.д.

Средний уровень выполнения конкретных операций фильтрации: гистограммирование, прослеживание, контроль, формирование выходных массивов, представление данных для выдачи рисунков и т.д.

Третий, высший уровень управления и организации работы основных блоков и общее управление программой.

В соответствии с этим строилась иерархия модулей программы FILTR, которая будет перечислена ниже. Отметим, что к модулям нижнего уровня отнесены многие достаточно сложные подпрограммы, которые могут вынуждать некоторые более простые подпрограммы или модули того же уровня, но они включены в этот уровень по своему положению в иерархии. Приведенные ниже в скобках названия подпрограмм даны как примеры, которые не исчерпывают, конечно, всех подпрограмм программы FILTR.

(1) Модули нижнего уровня

Манипуляции с памятью: чтение (RUPEOF,SKIP), чистка и пересылка массивов (UZERO,UCOPY), распаковка (UNPACK); формирование адресов и засылка массивов в Q; вычисление расстояний от точки до трека (DEVIAT); манипуляции с треками: открытие и закрытие трек-банка (OPENTR, CLOSET), вычеркивание треков (TRRJCT, KILL), пересылка треков в массиве TM при вычеркивании трека (MOVETR); гистограммирование (IHIST); поиск точек в коридоре вокруг предсказанного положения (SEARCH); подгонка м.н.к. (LSFIT); печать данных о рабочей памяти, протоколов фильтрации, рисунков и т.д.

(2) Функциональные модули второго уровня

Гистограммирование от вершины (HISTUP) и к вершине (HSTDWN); прослеживание от- и к вершине (FROMVR, UPTOVR), вызываемые из

общей подпрограммы прослеживания (TRACK) в зависимости от значения входного параметра; составление статистических гистограмм (HISTOG), подготовка и печать рисунков (PLHIST, PLRTH).

(3) Подпрограммы, организующие работу основных блоков (рис. 4):

Инициализации (READC, INITEV, INITVV), фильтрации (RECOGN, HUNT, BEAM), отсева ложных треков (SWOOP, TMSELN) сборки событий (OUTFIL).

Как уже отмечалось, общее управление осуществляет подпрограмма LEADER.

Примерами высокой утилизации модулей нижнего уровня могут служить такие подпрограммы, как UCOPY(15), LSFIT(11), RJPNT(9), RDIGF(9), TMRPNT(8), DEVIAT(8) (в скобках указано, в скольких других подпрограммах вызывается данный модуль).

Всего в память ЭВМ CDC-6500 загружается 179 подпрограмм (из них 61 системная), занимающих вместе с потребной рабочей памятью около 30К слов.

В ходе автономной отладки основных блоков была показана средняя скорость фильтрации порядка 8-10 с. на событие. Однако этот показатель сильно зависит от множественности и зашумленности событий, от числа ОПТ и т.д. и должен быть уточнен при эксплуатации программы. Для его улучшения полезно выявить наиболее часто вызываемые подпрограммы для их перевода в автокод. В настоящее время число несистемных автокодных программ составляет около 10%, и мы сознательно не повышаем этот показатель, чтобы сохранить высокий уровень мобильности программы FILTR (90%) с целью более простого ее перевода на ЭВМ БЭСМ-6.

Для повышения надежности работы программы использовались перечисленные выше средства контроля входных данных, позволяющие исключать дефектные проекции, могущие привести к выбросу задачи из счета, а также активно применялись системные средства, обеспечивающие сохранность результатов фильтрации при выбросе задачи оператором ЭВМ или при исчерпании ресурсов ЭВМ, заказанных для обработки заданной партии событий.

Быстрому обнаружению причин ошибок или отбраковки событий программой фильтрации помогает система диагностических сообщений, а также сведений о рабочей памяти, протоколов каждого шага процесса фильтрации и представление в виде рисунков исходных данных и различного рода гистограмм, содержащих статистическую информацию об обрабатываемом материале.

Вывод этих данных может быть произведен по желанию пользователя по специальным ключам печати.

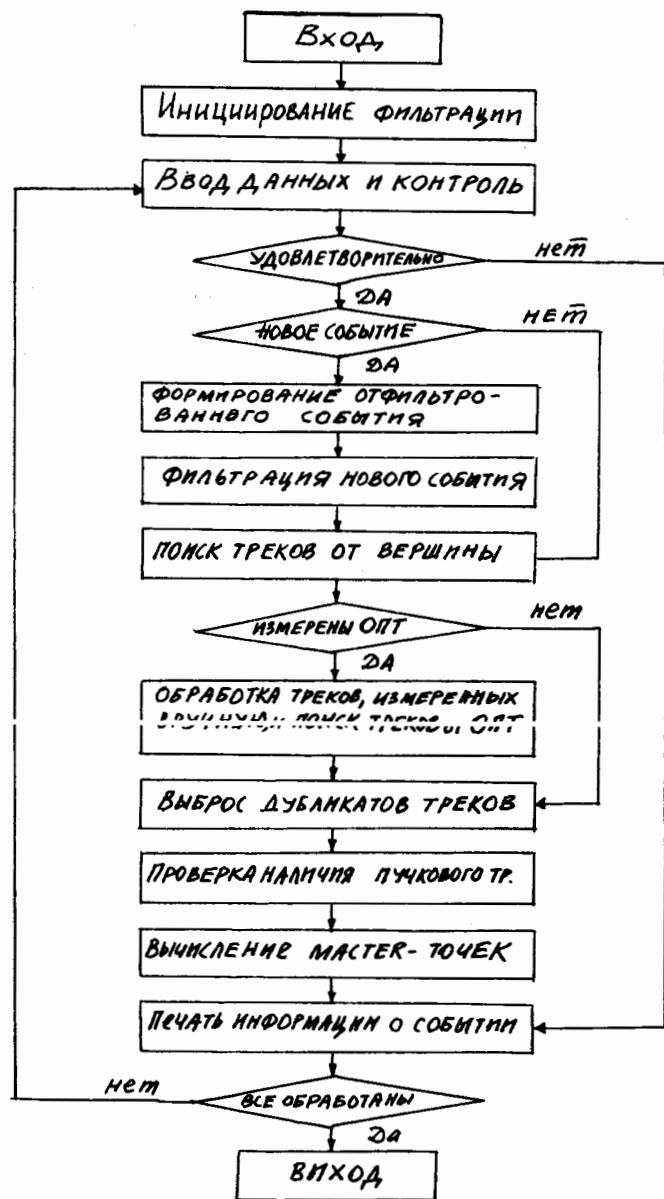


Рис. 4

Литература

1. Астахов А.Я. и др. ОИЯИ, РЮ-4943, Дубна, 1970.
Котов В.М. и др. ОИЯИ, Ю-7939, Дубна, 1974.
2. Lynch G.R. Preprint UCRL-17328, Univ. of California, 1967.
3. Pichon G. in: Proc. Eupean Spiral Reader Symp., Stockholm, 1972, CERN 72-16, pp.171-190.
4. Закамский Л.Л. и др. Препринт ИФВЭ, ОЭА 78-126, Серпухов, 1978.
5. Странд Р. В сб.: Распознавание образов при помощи цифровых вычислительных машин. Пер. с англ. "Мир", М., 1974, с.15.
6. Ососков Г.А. В кн.: Материалы семинара по обработке физической информации. Изд-во ЕрФИ, Ереван, 1975, с.268.
7. Готояия Т.Н. и др. В кн.: Материалы II Всесоюзного семинара по обработке физической информации. Изд-во ЕрФИ, Ереван, 1978, с.308.
8. Нергуй Б. ОИЯИ, РЮ-80-302, Дубна, 1980.
9. Косарева З.М. и др. В кн.: Всесоюзное совещание по автоматизации научных исследований в ядерной физике. Изд-во АН КазССР, Алма-Ата, 1978, с.153.
10. Ососков Г.А., Комолова В.Е. ОИЯИ, Ю-11290, Дубна, 1978.
11. Абдурахимов А.У. и др. ОИЯИ, Ю-80-831, Дубна, 1980.
12. Косарева З.М. и др. ОИЯИ, Ю-5574, Дубна, 1971.
13. Косарева З.М. и др. ОИЯИ, РЮ-8613, Дубна, 1975.
14. Ососков Г.А. ОИЯИ, РЮ-83-187, Дубна, 1983.
15. Ососков Г.А. ОИЯИ, РИ-4463, Дубна, 1969.
16. Абдурахимов А.У. ОИЯИ, Ю-81-359, Дубна, 1981.
17. Нергуй Б., Ососков Г.А. ОИЯИ, Б1-Ю-Ю370, Дубна, 1977.
18. Böck R. et al. В кн.: Труды международного симпозиума по вопросам автоматизации обработки данных с пузырьковых и искровых камер. ОИЯИ, ДЮ-6142, Дубна, 1971, с.547.
19. Дорж Л. и др. ОИЯИ, Ю-6882, Дубна, 1973.
20. UPDATE, Reference Manual, CDC Publication No. O449900, Minnesota, 1970.

Рукопись поступила в издательский отдел
25 августа 1983 года

Нергуй Б., Ососков Г.А.

P10-83-612

Основные алгоритмы и структура программы FILTR

Излагаются математические методы фильтрации и принципы организации программы FILTR, предназначенной для начальной обработки данных, полученных с помощью спирального измерителя при сканировании стереоснимков $\alpha - p$ взаимодействий, экспонированных на 1-метровой водородной камере. Предложена структура данных, оптимизирующая работу программы по скорости и требуемой памяти ЭВМ. Модульная архитектура программы повышает ее надежность, мобильность, упрощает процессы настройки параметров фильтрации и модернизации программы. Приведены основные характеристики программы.

Работа выполнена в Лаборатории вычислительной техники и автоматизации ОИЯИ.

Сообщение Объединенного института ядерных исследований. Дубна 1983

Nergui B., Ososkov G.A.

P10-83-612

The Main Algorithms and Structure of the FILTR Program

Mathematical methods of data filtering and principles of the FILTR program organization are considered. FILTR is intended for handling the data taken by JINR spiral reader from stereofilms of $\alpha - p$ interactions exposed on 1m hydrogen bubble chamber. The suggested data structure optimizes the program speed and needed computer memory. The modular architecture of the program increases its reliability and mobility. It simplifies processes of a filtering parameter adjustment and the program modernization. The main characteristics of the program are presented.

The investigation has been performed at the Laboratory of Computing Techniques and Automation, JINR.

Communication of the Joint Institute for Nuclear Research. Dubna 1983

Перевод О.С.Виноградовой