

Ц 848

0-754



ЛЕКЦИИ
ДЛЯ МОЛОДЫХ
УЧЕНЫХ

Г. А. Ососков

Применение
методов распознавания образов
в физике высоких энергий

ДУБНА

3111 ч. 102

ЛЕКЦИИ ДЛЯ МОЛОДЫХ УЧЕНЫХ

Выпуск 26

РЕДАКЦИОННЫЙ СОВЕТ

Д.В.Ширков - председатель
А.Н.Сисакян - зам. председателя
Г.М.Гавриленко - ученый секретарь
В.Г.Маханьков
В.А.Никитин
В.Р.Саранцева
С.Ю.Шмаков

© Обь

14017 в р.

P10-83-187

Г.А.Ососков

Ц 845
0-754

ПРИМЕНЕНИЕ
МЕТОДОВ РАСПОЗНАВАНИЯ ОБРАЗОВ
В ФИЗИКЕ ВЫСОКИХ ЭНЕРГИЙ

Объединенный институт
теоретической и экспериментальной
физики

Дубна 1983

Ц 8481

Ососков Г.А.
Применение методов распознавания образов
в физике высоких энергий

P10-83-187

После краткого введения в теорию и методы распознавания образов формулируется общая стратегия для фильмовых и бесфильмовых методов съема данных. Излагаются основные способы фильтрации и сжатия информации с трековых камер. В заключение обсуждаются перспективы развития методов обработки экспериментальной информации в физике высоких энергий.

Работа выполнена в Лаборатории вычислительной техники и автоматизации ОИЯИ.

Ososkov G.A.
Pattern Recognition Applications in High Energy Physics

P10-83-187

After the short introduction to the theory and methods of the pattern recognition a general strategie of data handling is formulated for the film and filmless data taking. The main methods of track chamber data filtering and compression are considered. Concluding some modern trends of the development of experimental data handling in high energy physics are discussed.

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

Перевод автора.

I. Введение

Взаимообусловленное влияние быстрого развития вычислительной техники на прогресс в области фундаментальных наук неизбежно ведет к все расширяющейся автоматизации научных исследований.

Одним из ярких примеров этого процесса может служить физика высоких энергий, где высокостатистические эксперименты, ставшие возможными благодаря внедрению достижений электроники в экспериментальные установки, порождают такой огромный поток опытных данных, что сложнейшая переработка его для извлечения желаемых физических результатов немислима старыми методами без привлечения средств автоматизации с использованием ЭВМ уже на самых первых стадиях отбора и обработки данных.

ЭВМ успешно применяются для проведения геометрического и кинематического анализа измеренных траекторий элементарных частиц. Однако в деле отбора информации для анализа на машинах образовался разрыв между высокой производительностью современных ЭВМ, выполняющих миллионы операций в секунду, и примитивной техникой подготовки данных для ввода в ЭВМ. Устранение этого несоответствия в массовой обработке стереофотографий событий в пузырьковых камерах путем введения сканирующих устройств для автоматического обмера этих снимков поставило еще одну сложную проблему. Заменяв человека в медленно выполнявшейся им функции измерений, ЭВМ оказалась вынужденной взять на себя также и такие операции (которые человек раньше выполнял без труда по ходу измерений), как опознавание событий на снимке, поиск центра взаимодействия и нахождение следов частиц, выходящих из центра. Разрешение этой проблемы пришлось искать на пути использования методов распозна-

вания образов, причем специфика и сложность этой задачи стимулировали известные пионерские работы Р.Нарасимхана в этой области по лингвистическому подходу к распознаванию. Хотя предложенная им реализация этого подхода потребовала использования такой супер-ЭВМ, как ИЛЛИАК-III^{3/}, сама проблема успешно решена обычными способами в физических центрах, в частности, в ОИЯИ, благодаря:

1) отказу от абстрактных грамматических правил описания сложной иерархии признаков трехмерного события и замены их на правила, основанные на физических ограничениях; 2) правильно выбранной стратегии обработки этих чрезвычайно зашумленных изображений, основанной на максимальном редуцировании шумов на возможно более ранней стадии поиска и оптимальному включению в него человека. Это позволило свести задачу трехмерного поиска к фильтрации одномерных элементов на двумерных проекциях. С математической точки зрения такое элементарное изображение, как трек, после ввода в ЭВМ может быть описано как дискретная последовательность точек, связанных близостью к гладкой кривой - траектории частицы в магнитном поле. Отклонения точек от кривой распределены по нормальному закону. Помимо параметров кривой, определению подлежит также и плотность распределения точек трека вдоль нее (так называемая ионизационная плотность). В настоящей лекции после краткого введения в теорию и методы распознавания образов формулируется общая стратегия обработки для фильмовых и бесфильмовых методов съема данных, излагаются основные способы прослеживания треков, методы сжатия информации с трековых камер и обсуждаются перспективы развития методов обработки экспериментальной информации в физике высоких энергий.

2. Основные этапы и методы распознавания образов

Первый этап связан с представлением исходных данных. Автоматизированная система обработки изображений должна включать сканирующее устройство, предназначенное для ввода оптической информации в ЭВМ. При этом изображение оцифровывается так, что каждой точке плоскости изображения с двумя координатами x и y ставится в соответствие функция $v(x, y)$, характеризующая степень почернения или яркость в этой точке. Набор всех таких дискретных точек или ячеек образует поле зрения, называемое растром, размерность которого определяет разрешающую способность нашей сис-

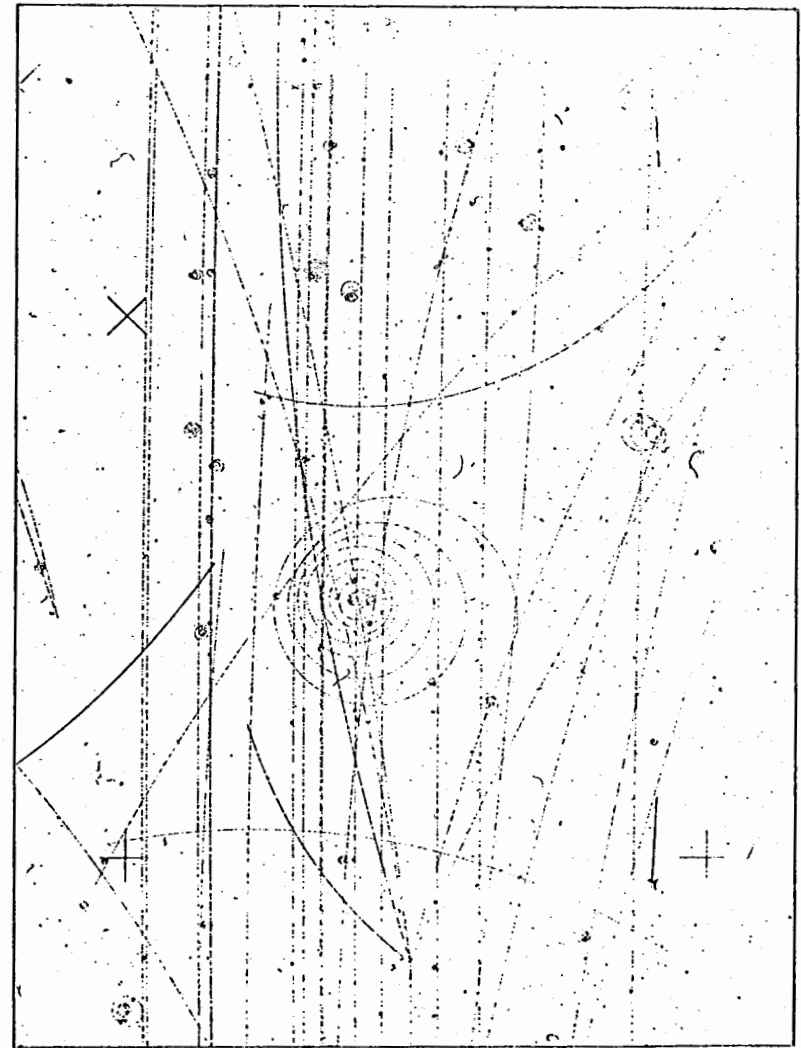


Рис. I.

темы. Количество значений, принимаемых функцией $v(x, y)$ и зависящих от чувствительности системы, обычно является степенью двойки и не превышает $2^5 + 2^6$. Если функция $v(x, y)$ принимает только два значения 1 и 0, то говорят о черно-белом изображении,

если значений больше двух, то изображение называется полутоновым. Наиболее характерным для изображений в физике высоких энергий является то, что они по большей части являются черно-белыми и носят штриховой характер, т.е. количество нулей на них может на несколько порядков превосходить число единиц, а точки, соответствующие единицам, образуют связные множества (см. рис. 1). Поэтому такие изображения более экономно (по сравнению с кодом Хэмминга, например) кодировать в виде $2n$ -мерного вектора $(x_1, y_1, x_2, y_2, \dots, x_n, y_n)$, где n — число черных точек изображения, а x_i, y_j — их координаты.

Второй этап распознавания образов связан с анализом исходных данных для выделения некоторых абстрактных свойств или признаков, присущих каким-то классам изображений и позволяющих определить их как некоторые образы.

Признаки несут меньше информации, чем исходное изображение, и поэтому служат средством сжатия этой информации, но они не должны терять тех существенных свойств, по которым на следующем этапе производится идентификация и классификация, т.е. отнесение данного изображения к одному из классов. Этот процесс и называется распознаванием.

Осуществляется он на третьем этапе путем применения к множеству признаков решающего правила.

В общей постановке можно рассматривать признак как действительное число, получаемое в результате применения к изображению, т.е. к функции $v(x, y)$ некоторого функционала $F(v(x, y))$. Набор k таких действительных чисел можно рассматривать как точку в k -мерном пространстве. Принятие решения о принадлежности каждой точки к тому или иному классу может быть выполнено, если есть какие-либо сведения о границах областей, занимаемых классами в нашем пространстве, или если в нем для каждого класса задано вероятностное распределение. Тогда, исходя из того или иного понятия геометрической близости или максимальной вероятности, мы можем осуществить распознавание, отнеся данное изображение к тому классу, в область которого попала точка признаков.

К сожалению, обычно на практике границы областей или распределения вероятностей неизвестны, и эти сведения можно получить только из эксперимента.

Появляется проблема обучения распознаванию, состоящая в том, что задается обучающая выборка, т.е. набор изображений с указанием для каждого из них соответствующего ему класса, по которой и ищутся границы областей. Таким образом в результате обучения строится такое решающее правило, которое обеспечивает правильную классификацию не только изображений из обучающей выборки, но и всех последующих, предъявляемых для опознавания.

Правильность здесь понимается в смысле минимальности ошибки распознавания, которая возможна при данной длине обучающей выборки.

Задача, в которой указание классов отсутствует, а требуется выделить классы изображений, похожих с точки зрения некоторого критерия близости или статистических свойств описаний этих классов, называется задачей распознавания без обучения.

В зависимости от априорных сведений о классах изображений методы распознавания основываются на предположениях либо о типах решающих функций, т.е. функций, задающих разбиения множества всех значений признаков (непараметрические методы), либо о вероятностных свойствах признаков, задаваемых условными распределениями, параметры которых и оцениваются в процессе обучения (параметрические методы). Задачи без обучения решаются методами кластеризации, т.е. объединения точек в пространстве признаков в группы, — кластеры, — по принципу близости.

Примером первой группы непараметрических методов может служить метод потенциальных функций, когда каждой точке из k -го класса обучающей последовательности приписывается потенциальная функция, максимальная в этой точке и убывающая с расстоянием от нее. При опознавании нового изображения его относят к тому классу, суммарный потенциал которого в точке, соответствующей этому изображению, больше потенциалов других классов. В^{35/} описаны результаты применения этого алгоритма для идентификации частиц высоких энергий, проходящих через телескоп из четырех детекторов.

В качестве иллюстрации ко второй группе параметрических методов приведем более сложный пример распознавания незрительных образов из физики высоких энергий, связанный с использованием искровых или проволочных камер на линии с ЭВМ, позволяющих оцифровать траектории частиц, минуя стадию фотографирования и последующих измерений. $3n$ -мерный вектор, описывающий пространственную траекторию заряженной частицы в магнитном поле, давшую n отсчетов в проволочных камерах, зависит всего от 5 физических параметров. Поэтому такие векторы, соответствующие

реальным траекториям, должны группироваться вблизи некоторой пятимерной гиперплоскости. Поскольку связность точек в траектории можно характеризовать корреляционной матрицей V_x , то естественно искать такое новое пространство, в котором V_x приняла бы диагональный вид. Таким пространством, как известно, является пространство собственных векторов матрицы V_x , причем главные члены этого разложения по собственным значениям V_x , называемого разложением Карунен-Лова^{/36/}, и должны описывать нашу пятимерную плоскость и ее ближайшую окрестность. Линейное преобразование в это новое пространство с размерностью, значительно меньшей, чем исходное, и означает переход к наиболее информативным признакам ξ_k . Из-за наличия экспериментальных ошибок фактическое число признаков должно быть больше пяти, с тем, чтобы, считая

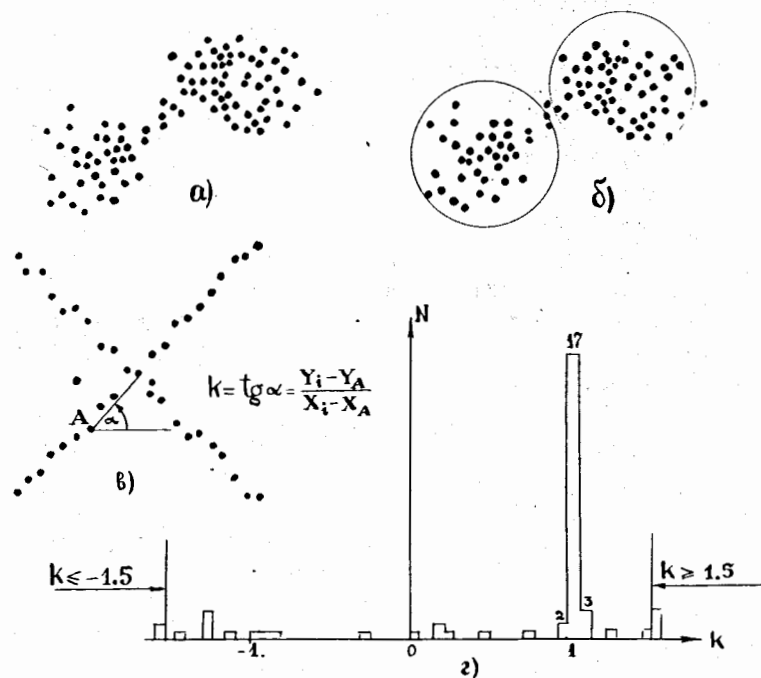


Рис.2.

главными первые 5 компонент, использовать остальные для проверки близости к 5-мерной плоскости, т.е. в качестве решающей функции берется сумма $d^2 = \xi_6^2 + \xi_7^2 + \dots$. Следует отметить важные проблемы, составляющие главные, часто непреодолимые, трудности в подобных задачах распознавания:

- необозримость множества комбинаций, возникающих при попытке объединить отсчеты искровых камер в многомерные векторы, соответствующие траекториям;
- громоздкость процедуры проверки граничных условий в многомерном пространстве;
- наличие пропусков отсчетов на траектории частицы, вызванных несрабатыванием той или иной камеры, что приводит к исключению вектора из рассматриваемого пространства.

Успех применения метода главных компонент для обработки данных спектрометра R603 ^{/37/} объясняется именно тем, что удалось преодолеть все три перечисленных фактора. Первый - благодаря быстрой процедуре преселекции комбинаций, основанной на точном учете геометрии установки, второй - с помощью перехода в пространстве главных компонент и третий - в силу высокой эффективности камер и установки R603 и однородности магнитного поля.

Однако, пожалуй, чаще всего при обработке изображений в физике высоких энергий используются методы третьей группы - методы кластеризации^{/38/}.

При этом, как и в первых двух группах методов, чрезвычайно важным является вопрос о правильном выборе признаков. На рис.2 даны два примера кластеров точек. В случае 2а в качестве признака сходства двух образов достаточно взять их взаимное расстояние. Вычислив расстояния между всеми парами точек и применив простейшую решающую функцию, - сравнение с выбранным порогом, - можно довольно быстро найти подходящий порог, который позволит разделить точки этих двух кластеров, как показано на рис.2б. Однако, если различные образы пересекаются, как на рис.2в, то предложенная пороговая процедура окажется бессильной, пока мы не введем еще один признак, характеризующий принадлежность точек к "своему" кластеру. Этот признак - линейная связность точек, которую можно количественно выразить как близость угловых коэффициентов прямых, соединяющих точки одного и того же кластера. Вторым признаком более весом, чем первый, и в простых случаях, когда нет шумов и все точки могут быть отнесены либо к одному, либо к другому -

му классу, можно обойтись только этим вторым признаком. Возьмем произвольную точку А, вычислим угловые коэффициенты всех отрезков, соединяющих эту точку с остальными точками на рис.2в, и занесем их абсолютные величины в гистограмму, представленную на рис.2г. Точки, давшие пик в этой гистограмме, можно с большой уверенностью отнести к кластеру, включающему точку А. При наличии шумовых точек, не относящихся ни к одному из двух кластеров на рис.2в, необходимо использовать оба признака. Отметим, что образы, подобные тому, который изображен на рис.2в, возникают при сканировании реперного креста - одной из самых массовых фигур на снимках с трековых камер. Реперные кресты необходимы для восстановления пространственной картины события и для учета возможных дисторсий, поэтому их центры подлежат особо точному измерению.

Распознавание крестов среди других данных сканирования облегчается тем, что приблизительно известен район их расположения и задана форма: угол между плечами креста и их длина. При наличии большого количества шумовых отсчетов эти априорные данные позволяют с большим успехом использовать параметрические методы. В одном из них^{1/} пара прямых, образующих крест, описывается выведенным уравнением кривой второго порядка, зависящим от трех параметров (координат центра и угла поворота всего креста относительно заданного угла), которые оцениваются методом наименьших квадратов по всем точкам.

В случаях, когда изображение очень сложно и характерных признаков слишком много или их не удается идентифицировать, можно на основе анализа структуры образа выделить некоторые элементарные изображения, складывающиеся в более сложные подобразы, которые могут образовать целую иерархическую систему. Нам требуется набор особых правил, чтобы описать свойства и взаимное расположение подобразов. Нетрудно усмотреть здесь аналогию с синтаксической структурой языков, поэтому такой подход в распознавании образов называется лингвистическим^{2/}, а процедура описания - структурным анализом изображения.

Как уже упоминалось, одной из первых работ по лингвистическому методу была статья Р.Нарасимхана^{3/}, который занимался распознаванием событий на снимках с пузырьковых камер. Однако следует отметить, что знакомство с примерами обработки, приведенными в^{3/}, показывает, что даже с помощью полно сформулированной грамматики проблему ликвидации несоответствий в работе программы

BUBBLE SCAN, обусловленных разрывами треков и неправильным их соединением, можно решить лишь частично путем учета некоторых априорных сведений о рассматриваемых классах изображений. Как неоднократно подчеркивал сам Р.Нарасимхан, неперенными условиями применения лингвистического подхода являются: а) использование сверхбыстрых ЭВМ (десятки млн. операций в секунду) с параллельной обработкой информации, таких, как ИЛЛИАК-Ш; б) предварительная обработка оцифрованной информации, подлежащей структурному анализу для заполнения разрывов и утоньшения линий. Кроме этого, программа BUBBLE SCAN работает в интерактивном режиме, т.е. с участием человека, корректирующего работу ЭВМ.

3. Распознавание образов при обработке данных трековых камер

Чтобы понять, как же все-таки обрабатываются такие сложные изображения в физических центрах, из которых даже самые крупные не имеют столь мощных вычислителей как ИЛЛИАК-Ш, следует начать с более подробного знакомства с изображениями событий в трековых камерах.

Взаимодействия, происходящие в таких камерах, фотографируются, по крайней мере, тремя фотообъективами (их число может достигать и до восьми), чтобы "увидеть" все рабочее пространство камеры и обеспечить правильность пространственной реконструкции. На каждой проекции мы имеем изображения одного или нескольких событий множественного рождения частиц, состоящие из треков, один из которых, называемый первичным, или "пучковым", оставлен частицей, претерпевающей взаимодействие в мишени, и порождает большое количество вторичных треков. На фотографиях событий с пузырьковых камер, мишенью в которых является сама жидкость, заполняющая камеру, человек легко обнаруживает проекцию центра взаимодействия как место, называемое вершиной, где один из пучковых следов разветвляется в характерную метелку вторичных треков с направлениями, зависящими от энергии пучка и типа произошедшего взаимодействия (рис.1). Треки состоят из цепочек пузырьков со средним диаметром 15+30 мкм. Кроме этого, на фотографии с пузырьковой камеры имеется служебная информация (реперные кресты, коды номеров кадров и т.д.), а также масса фоновой информации в виде не взаимодействовавших пучковых и просто посторонних треков, отдельных пузырьков, царапин, бликов от системы освещения и т.д. Эта фоновая информация обуславливает необычайно высокую зашумленность

фотографий с пузырьковых камер. Например, после работы измерительного автомата с параллельной разверткой в память ЭВМ с одной только фотографии может поступить свыше 60 тыс. пар координат, из которых не более 4+5 тыс. принадлежат трекам события.

Отметим при этом, что точность измерения центров пузырьков не должна быть ниже 10% их диаметра, что для двухметровой камеры означает точность 10^{-5} . Некоторое повышение точности достигается путем усреднения 10-30 последовательных отсчетов от одного трека в так называемые мастер-точки.

Такие высокие требования в сочетании с большой зашумленностью и объясняют неприменимость классических методов распознавания образов. Собственно задача обработки камерных снимков по своей природе не является обычной задачей классификации объектов. Обнаружению подлежат не треки, а более сложные пространственные объекты - события, которые восстанавливаются затем в ходе геометрической реконструкции по своим отдельным элементам-трекам, которые должны быть связаны между собой, так как являются проекциями следов частиц, встречающихся в одной точке пространства, соответствующей вершине события.

Таким образом, в самой постановке задачи заложено разбиение сложного объекта на иерархию фрагментов, распознавание которых на самом низком уровне - проекций треков - является более простым делом, так как они представляют некоторые линейные образы на плоскости. Существующие программы геометрического и кинематического анализа информации камерных экспериментов для извлечения из нее физических параметров рассчитаны именно на такой вход, в котором информация о событии подается в виде последовательных списков координат мастер-точек всех треков по проекциям события. Эти списки данных, включающие для пузырьковых камер также и координаты вершин, возможно, и вторичных вершин и выходящих из них треков (если вторичная частица также претерпевает взаимодействие) точек изломов, остановок треков и т.д., являются примером подробных описаний событий на специальном многоуровневом языке, разработанном физиками и программистами.

Можно сразу же отметить, что значительная роль в этом языке отведена описанию разного рода служебной информации, без которой невозможна правильная обработка данных о событии: координаты реперных крестов, калибровочные константы, сведения о параметрах камеры и т.д., но мы не будем здесь касаться этой стороны вопроса.

Важно, что если это описание удовлетворяет всем требованиям по точности определения дескрипторов, то оно позволяет затем восстановить в пространстве камеры все треки события и вычислить все необходимые физические величины. Таким образом, вместо чисто лингвистического подхода с использованием формальных грамматических правил, описывающих взаимное расположение элементов треков для объединения их в события, более успешным оказался подход с использованием наиболее естественных в этой обстановке геометрических и кинематических правил и формул, полнее учитывающих всю априорную информацию и специфику эксперимента.

Но не следует забывать, что успех дела с точки зрения создания автоматической системы распознавания был определен прежде всего на "элементарном" уровне распознавания треков события среди всех оцифрованных координат, поступающих в ЭВМ при сканировании изображения очередной проекции события.

Рассмотрение различных систем автоматической обработки камерных снимков показывает, что прогресс в этой области был достигнут, прежде всего, отказом от полной универсальности и выбором разумного компромисса между возможностями ЭВМ и способностями человека путем включения его в систему. Если пойти дальше, по пути анализа алгоритмов и программ обработки и выделения общих принципов, составляющих стратегию успеха их работы, то можно сформулировать следующие основные принципы этой стратегии.

1. Полезны лишь менее 10% информации на снимке, поэтому ненужную информацию следует отсеивать на возможно более ранней стадии обработки.

2. Долго и дорого обрабатывать все данные с одной и той же высокой точностью, отказ же от нее может привести к потере физической информации, поэтому необходимо максимально использовать априорные знания и увеличивать точность только в ходе дальнейшего поиска.

3. Простота алгоритма поиска может сделать его главным "растратчиком времени" в программе, поэтому следует избегать перебора или, по крайней мере, сокращать в каждом цикле множество точек, по которому ведется перебор; чем раньше поиск станет одномерным, тем лучше.

4. Система без участия человека всегда оказывается неоправданно дорогой и узкоспециализированной, поэтому включение челове-

ка в распознавание целесообразно, но оно не должно замедлять нормальную работу системы.

Б. Единственно надежный критерий качества распознавания - достоверность физических результатов, поэтому в системе должна быть быстрая обратная связь, включающая программы для непосредственного получения окончательных физических результатов обработки.

Несмотря на внешнюю очевидность каждого из этих принципов, неуклонное следование им всегда позволяло выиграть время и добиться успеха в решении компромисса между надежностью и скоростью. Очевидно, что эти принципы не противопоставлены общей схеме распознавания образов, а напротив, дополняют ее, учитывая специфику обрабатываемой информации.

В соответствии с этой стратегией и общей процедурой распознавания разобьем процесс нахождения и точного описания событий на следующую последовательность основных этапов, по которым и будет вестись дальнейшее изложение:

- автоматическая дискретизация стереоснимков;
 - предварительная обработка данных;
 - выделение элементов треков;
 - распознавание треков;
 - составление описания события для ввода в программы геометрического и кинематического анализа;
 - обучение системы обработки путем внесения коррекций в параметры решающих функций по результатам анализа ошибок распознавания.
- Отметим, что в эту схему укладываются и бесфильмовые методы обработки камерных данных.

4. Методы автоматической обработки фотографий с пузырьковых камер

4.1. Способы дискретизации изображений

В соответствии со способом сканирования, т.е. просмотра снимка световым лучом, сканирующие устройства можно классифицировать на ряд групп. К первой из них отнесем устройства с глобальным просмотром всего поля сканирования, они основаны на принципе механической развертки и отличаются только траекторией и видом светового пятна. Устройство Хафа и Пауэлла - нрд^{/4/} или аналогичные отечественные устройства^{/5/} осуществляет параллельную развертку пятном круглой формы. Спиральный измеритель сканирует изображение по спирали, радиально ориентированной узкой щелью от вершины события, измеряемой оператором^{/39/}.

Другая группа устройств использует электронно-лучевое сканирование, гибкость которого позволяет совмещать сканирование с поиском вершин и треков, осуществляя растровую развертку отдельных участков снимка. Сюда относятся установки REPR, POLLY, ERASME, МЭЛАС, ЭЛАС, АЭЛТ-П и др.

Можно еще упомянуть устройства типа SWEEPNIK, осуществляющие аппаратное слежение за треком с помощью щели, вращающейся вокруг одного из своих концов. Наведение щели на каждый из треков выполняется оператором. Аналогично работает система NEVAS^{/40/}

Подробные описания всех этих измерительных установок можно найти в обзорах^{/4,6/} и трудах международных конференций^{/7,26,62/}. Далее мы будем рассматривать вопросы обработки данных устройствами первой группы, так как задачи распознавания образов с участием человека, характерные для остальных устройств, выходят за рамки нашей лекции.

Малый размер пятна, обеспечивающий высокую точность нрд (~ 2 мкм), позволяет получить со снимка размером 100x55 мм до 60 тыс. координат, количество которых может удвоиться, если в событии окажутся треки, параллельные направлению сканирования, и придется прибегнуть к так называемому аномальному сканированию в перпендикулярном направлении. Для сокращения этого потока данных прибегают к методу поведения по дорожке, состоящему в использовании данных предварительных грубых измерений вершины и двух точек на каждом треке события для построения по этим данным дуговых дорожек шириной около 0,5 мм с тем, чтобы в процессе сканирования отбирать для запоминания координаты только тех точек, которые лежат внутри дорожки. Являясь примером полной реализации первого из вышеперечисленных принципов стратегии сканирования, метод наведения по дорожке, помимо кардинального сокращения количества данных, значительно упрощает поиск треков и повышает точность измерений. Однако этот метод требует значительных затрат времени и оборудования на предварительные ручные измерения, которые окупаются только в случаях событий сложной конфигурации.

Используется также и другой метод - наведение по вершине, существенно снижающий затраты на предварительные измерения, так как при этом вручную измеряются только координаты вершин событий и короткие треки, но требующий уже большой ЭВМ для работы на линии с нрд, чтобы во время сканирования осуществлять инициацию трековых элементов и поиск треков.

Имеются примеры работы НРД и без наведения (см. например, /8,34/, но эти системы ограничены определенным типом экспериментов, требуют высокой частоты появления событий на кадрах фильма и/или больших вычислительных средств.

Метод наведения по вершине используется также при работе спирального измерителя (СИ), где оператор, пользуясь декартовой системой координат, измеряет вершину события, концы коротких треков (КТ), а также специальные опорные точки (ОТ) на зашумленных треках. После этого в процессе спиральной развертки от вершины в ЭВМ передаются полярные координаты всех точек, пересекаемых этой спиралью, причем радиально ориентированная щель в районе вершины "видит" только треки, выходящие из нее или проходящие совсем близко. Фильтрующий эффект щели и меньшая плотность витков спирали в сравнении с частотой параллельной развертки приводит к тому, что весь массив измеренных полярных координат $(R_i, \theta_i)_{i=1,n}$, где n не превышает 4 тыс. и лежит обычно в пределах 1000-2000 точек, может быть записан на магнитную ленту для последующей (off-line) обработки. Последовательность полярных координат, относящихся к данной проекции, снабжается паспортом проекции, содержащим, помимо служебной информации (номера эксперимента, события, проекции, даты измерения и т.д.), декартовы координаты всех точек, измеренных в этой системе (вершины, кресты, КТ и ОТ).

4.2. Предварительная обработка данных

Цель этого этапа - не допустить к дальнейшей обработке данные, имеющие дефекты как в служебной информации, не используемой непосредственно для распознавания, так и в самих координатах, могущих это распознавание исказить. Несколько упрощая, представим себе, например, что одно из чисел в последовательности данных СИ (R_i, θ_i) будет пропущено при поблочной записи на магнитную ленту. Тогда во всем последующем блоке, т.е. в кольцевой зоне вокруг вершины, полярный радиус и угол поменяются местами. Нетрудно представить, как это может сказаться на распознавании треков! Подобную ошибку можно обнаружить, проверяя монотонность возрастания радиуса по всему набору $(R_i, \theta_i)_{i=1,n}$. Также должны быть проверены декартовы координаты реперных крестов и вершины. Поскольку позиции реперных крестов в камере точно известны, то после некоторых вычислений для перехода в соответ-

ствующую систему координат могут быть обнаружены недопустимые отклонения в измерениях реперных крестов, а также несоответствия между координатами вершины на разных проекциях.

Важность предварительного отсева недоброкачественной информации при массовой обработке на самой ранней ее стадии объясняется не только желанием избежать напрасных затрат времени ЭВМ на обработку, результаты которой все равно будут потом отвергнуты в последующих программах, но, что более серьезно, - это предохраняет нас от опасности искажения этими неверными данными, если они не будут отвергнуты, окончательных физических результатов. Кроме того, сведения об ошибках в данных, именно на этой стадии, позволяют очень хорошо контролировать работу самого сканирующего автомата.

Предварительная обработка может выполняться либо на малой ЭВМ, обслуживающей сканирующий автомат, либо, если эта ЭВМ не располагает нужными ресурсами скорости и памяти, на большой ЭВМ в качестве начального этапа off-line обработки.

К предварительной обработке может быть отнесена описанная в /9/ префильтрация, т.е. использование фильтрующего эффекта щели для быстрого обнаружения по первым виткам спирали начала треков, выходящих из вершины, и организации наведения по дорожке в виде грубого углового сектора. Это позволяет без потери темпа сканирования сократить поток данных в 4-5 раз. Аналогичная программа разработана в ОИЯИ /10/, но для использования в других целях.

Мы не затрагиваем здесь входящие в предварительную обработку, но не относящиеся к рассматриваемым нами вопросам, предварительные преобразования координат, необходимые для коррекции нелинейных искажений, вносимых измерительным устройством /11,41/.

4.3. Выделение элементов треков

С точки зрения применяемых методов задача выделения элементов треков, т.е. поиска в некоторой небольшой области групп точек, связанных в линейные цепочки, аналогична для приборов типа НРД или СИ, и может быть отнесена к кластерному анализу /38/.

Наиболее употребительным является метод поворотных гистограмм /4/, состоящий в разбиении исследуемой области на параллельные узкие полоски и подсчете числа точек, попавших в каждую из полос (рис.3). Постепенно меняя направление полосок, получим,

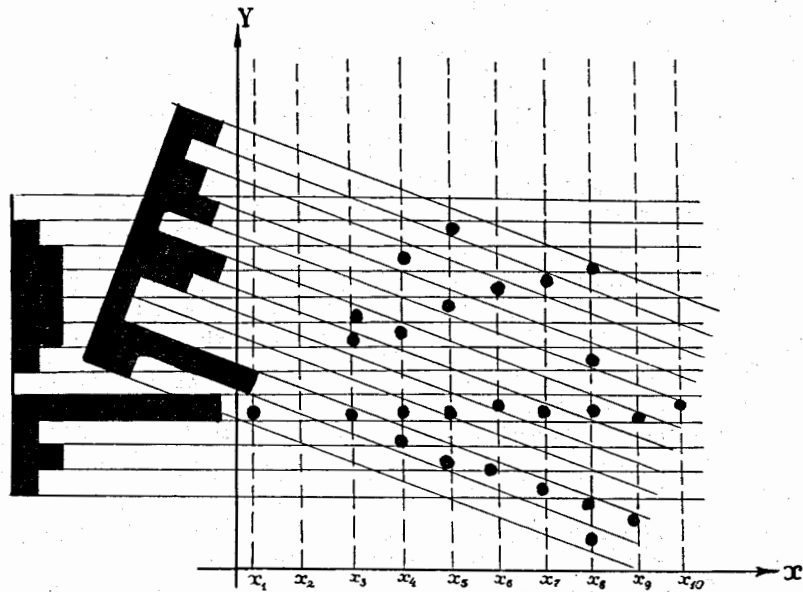


Рис.3.

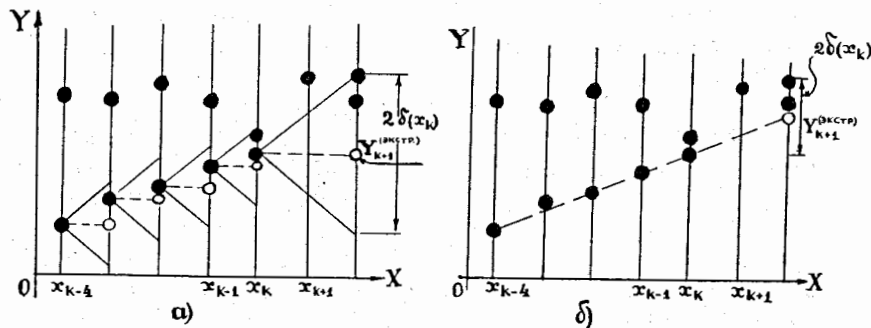


Рис.4.

что при совпадении его с направлением элементов трека точки этого элемента, попав в одну из полос, дадут высокий пик в гистограмме. Трек детектируется при превышении этим пиком порога, вычисляемого по треку с минимальной плотностью пузырьков. Длина каждой полоски обычно не превышает 10-12 скан-линий. Ширина полосок должна соответствовать средней ширине трека (20-30 мкм). Для того, чтобы не потерять треки, разрезаемые вдоль границы между полосками, они обычно перекрываются на 10-15% их ширины. Последнее обстоятельство, а также то, что трековый элемент может быть обнаружен еще дважды в двух соседних полосках при значениях углов поворота, предшествующих и следующих за истинным, приводит к появлению дубликатов, которые, однако, легко устраняются. Сложнее обстоит дело с "мнимыми" трековыми элементами, появляющимися на зашумленных участках из-за попадания в полоски отдельных шумовых точек или других треков, пересекающих область гистограммирования. Их устранение производится обычно на следующем этапе распознавания, как и анализ причин слишком высоких импульсов в гистограмме, которые, кроме наличия шумовых точек, могут свидетельствовать о попадании в одну полоску двух близко идущих треков.

Следующий, более быстрый способ обнаружения трековых элементов - метод шнуров (STRINGING) ^{6/} состоит в объединении точек в соответствии с их близостью по одной из координат. После объединения двух точек можно включить в критерий близости и некоторое направление, и по нему искать третью точку и т.д., при этом необходимо следить за линейностью полученного шнура, чтобы не уйти в сторону, присоединив к нему точку, не принадлежащую данному трековому элементу и случайно оказавшуюся близко. Поэтому длину шнура обычно ограничивают той же величиной - 10 точек, что и при гистограммировании (рис.4а).

Некоторое усовершенствование этого метода, состоящее в более обоснованной экстраполяции возможного положения следующей точки шнура, предложено в ^{12/}, где по шнуру из k точек с направлением, близким к оси OX, координаты (k+1)-й точки экстраполируются по рекуррентным формулам:

$$y_{k+1}^{(экстр.)} = y_k^{(экстр.)} + \frac{y_k^{(экстр.)}}{D_k G}; \quad x_{k+1}^{(экстр.)} = x_k + \Delta x. \quad (I)$$

Здесь (x_k, y_k) - координаты последней точки шнура,
 $(x_k^{(экстр.)}, y_k^{(экстр.)})$ - предыдущий прогноз, а информационный
 член D_k вычисляется как

$$D_k = q_k(x_k - x_{k-1}) + (1 - q_k)D_{k-1}, \quad (2)$$

q_k - убывающая функция k , Δx - шаг сканирования. Близость
 к прогнозируемой точке (I) понимается в смысле принадлежности
 y_{k+1} к интервалу ширины $2\delta(x_k)$ с центром в $y_{k+1}^{(экстр.)}$ ($\delta(x)$
 - убывающая функция x). При отсутствии точек в этом интервале
 увеличивается функция пропуска g в (I) и поиск переносится на
 следующую скан-линию. Если оказывается, что g превысила задан-
 ный порог, то шнур обрывается (рис.4б).

Очевидно, что в этом подходе уже нет необходимости ограни-
 чивать длину трекового элемента; собственно, благодаря процедуре
 экстраполяции стадия выделения трековых элементов объединилась
 со следующей - распознаванием треков. Мы вернемся к методу шнуров
 в следующем разделе.

Идея об объединении точек в трековый элемент на основании
 некоторого критерия близости была положена в [13] в основу алго-
 ритма, выделяющего сразу все трековые элементы, близкие к верши-
 не. Это осуществляется известным из теории графов методом наимень-
шего связующего дерева. Близость точек в радиальном направлении
 учитывалась с помощью специальной метрики, "предпочитающей" это
 направление. Поскольку метод применялся к данным спирального скан-
 нирования в некотором круге R с центром в вершине, то переход
 к такой метрике просто означал сжатие плоскости по полярному ра-
 диусу R .

Представление полярных координат r, θ в виде прямоугольной
 системы координат (рис.5,6) весьма полезно, так как позволяет
 применить для реализации метода поворотных гистограмм технику
 проектирования, значительно ускоряющую этот метод, самый надеж-
 ный, но и самый медленный из вышеперечисленных. Вместо многократ-
 ного перебора всех точек из области гистограммирования для про-
 верки их принадлежности к интервалу гистограммы (вспомним третий
 принцип) в методе проектирования вычисляются проекции точек по
 соответствующим направлениям. Для точки с координатами (r, θ)
 результат проектирования на полярную ось под углом α равен

$$e_{пр} = e - tg\alpha \cdot r. \quad (18)$$

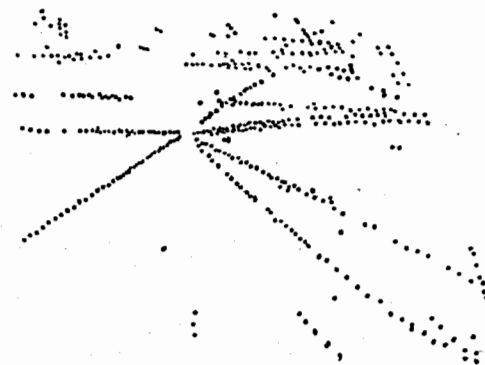


Рис.5.



Рис.6.

Меняя направление проектирования от α_{\min} до α_{\max} (см.рис.7) с шагом
 $\Delta\theta$, получаем, что для k -го направления $tg\alpha_k$ вычисляется
 как

$$tg\alpha_k = tg\alpha_{\min} + k\Delta\theta.$$

Превращая $e_{пр}$ в адрес, по которому в соответствующую гистограм-
 му будет добавлена единица, и меняя k , мы фактически находим вклад
 каждой точки во все гистограммы. Таким образом, для каждой ячейки
 $e_i = e_{i-1} - e_{i-1}$ на оси находится гистограмма, соответствующая вееру
 направлений, выходящих из этой ячейки.

Если полярный угол θ_0 , под которым трек выходит из вершины,
 таков, что $e_{i-1} \theta_0 e_i$, то этот трек генерирует пик в гистограм-
 ме в направлении, соответствующем направлению этого трека. В работе
 [42] дан быстрый алгоритм метода проектирования, позволяющий реализо-
 вать его аппаратно или выполнить в виде программы, работающей на ли-
 нии со спиральным измерителем. Для приборов с параллельным сканирова-
 нием типа НРД известны несколько успешно работающих устройств,
 таких, как BRUSH [77], СНР [29] или ПРЭТ [31], которые

позволяют аппаратно решить противоречие между высокой эффективностью метода поворотных гистограмм и требованием существенных ресурсов ЭВМ на его быстрое выполнение.

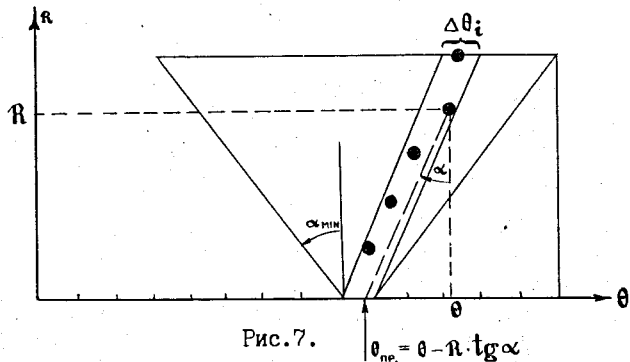


Рис. 7.

4.4. Распознавание треков и событий

Как уже было сказано, события на снимках с пузырьковых камер представляют собой образы с хорошо известной структурой, так что их распознавание сводится к синтаксическому управляемому анализу. Напомним структуру события:

- трековые элементы образуют трек;
- пересечение нескольких треков в одной точке определяет вершину;
- треки и вершины образуют двумерную проекцию события;
- три проекции вместе с реперными крестами образуют трехмерное событие.

Этой структуре соответствует и процедура распознавания.

При отсутствии априорных сведений о событии можно проводить воссоединение треков, связывая с каждым из трековых элементов пару: направление d и мастер-точку p (центр тяжести трекового элемента) и, начиная с какой-то иницилирующей пары (d_0, p_0) , искать в направлении d_0 следующий элемент (d_1, p_1) , ближайший к предыдущему по направлению. Пару элементов можно использовать для лучшей экстраполяции (уже не обязательно линейной) и т.д. аналогично тому, как это было в методе шнуров.

Именно так происходит поиск пучковых треков при работе без наведения. Затем все точки, где пучковые треки обрываются или меняют направление, сопоставляются по проекциям для определения вершин, после чего задача сводится к распознаванию с наведением по вершине.

Ниже мы еще вернемся к глобальным методам воссоединения трековых элементов в разделе об обработке информации при бесфильмовом съеме данных со стримерных камер, а здесь рассмотрим более типичную для обработки снимков с пузырьковых камер задачу с наведением по вершине.

Экстраполяция одного из трековых элементов, найденных вблизи вершины, на следующую скан-линию для поиска и присоединения к элементу еще одной точки, называется прослеживанием трека.

Существует объективная необходимость в такой операции поэлементного наращивания трека, но с обязательным использованием информации об уже найденных точках, вызванная наличием разрывов из-за слабого контраста пузырьков, прохождения трека через обширные пятна и пересечением с другими треками. Последнее особенно неприятно в случаях пересечения треков под малыми углами, так как любое из устройств детектирования середины сигнала при пересечении трека пятном имеет конечную разрешающую способность, что ведет к возникновению некоторого искаженного усредненного сигнала в зоне пересечения треков.

Длина этого отрезка, вносящего искажения в положение точек каждого из перескакивающих треков, тем больше, чем меньше угол пересечения. Исключение искаженных точек из массива трека не представляет проблемы, благодаря процедуре выброса, о которой речь пойдет ниже, но более опасно "свернуть" при прослеживании через место пересечения на ложное направление, выбрав неправильное продолжение.

В таких случаях все зависит от того, насколько учтена в экстраполяционной формуле информация, имеющаяся в уже найденных точках. Выделим два основных способа экстраполяции треков:

1. Проведение экстраполяционной кривой по методу наименьших квадратов (м.н.к.).
2. Предсказание байесовского типа.

Для выбора подходящей кривой, наилучшим образом описывающей трек, следует учесть, что движение заряженной частицы в однородном магнитном поле при отсутствии энергетических потерь происходит по винтовой линии, проекция которой на плоскость может быть хорошо приближена окружностью с радиусом, обратно пропорциональным импульсу частицы. Поэтому в задачах прослеживания треков использование уравнения окружности стало традиционным. Возвращаясь опять к случаю спирального сканирования, выведем полярное уравнение окружности радиуса r , с центром в точке

(R_c, θ_c), проходящей на малом расстоянии β от полюса так, что $R_c = r + \beta$. По известной тригонометрической формуле

$$R^2 + \beta^2 + 2r\beta + 2R(r+\beta) \cos(\theta - \theta_0) = 0,$$

откуда при β , малом по сравнению с радиусом окружности r , пренебрегаем членами β/r и β^2 и получаем

$$\theta = \theta_0 + \arccos\left(-\frac{R}{2r} + \frac{\beta}{R}\right).$$

Разлагая в ряд \arccos , получим известную аппроксимацию для уравнения окружности в полярных координатах в виде

$$\theta = \hat{\theta}_0 + \alpha R + \beta/R + \gamma R^3, \quad (3)$$

где $\alpha = \frac{1}{2r}$ - половина кривизны окружности, $\hat{\theta}_0 = \frac{\pi}{2} - \theta_0$, а $\gamma = \alpha^3/6r$.

Мы, однако, не будем пользоваться связью α и γ , чтобы сохранить линейную зависимость от коэффициентов. Это позволит нам найти их по имеющимся n точкам трека (R_i, θ_i) $i = \overline{1, n}$, минимизируя квадратичный функционал

$$s^2 = \sum_{i=1}^n w_i (\theta_i - \theta_0 - \alpha R_i - \beta/R_i - \gamma R_i^3)^2. \quad (4)$$

Уклонения точек (R_i, θ_i) от кривой лучше измерять не по углу θ , точность которого убывает с радиусом, а по дуге $l = R \cdot \theta$, т.е. сохранить декартову метрику. Для этого надо выбрать веса w_i пропорциональными R_i^2 . Приравняв нулю производные s^2 по параметрам, мы получаем относительно них нормальную систему линейных уравнений с симметричной матрицей. Решение этой системы дает нам оценки для θ_0, α, β и γ , а значение s^2_{\min} - оценку дисперсии, необходимую для организации процедуры выброса "плохих" точек.

Потребность в этой процедуре возникает, если средний разброс точек σ_n вокруг кривой (3) превысит допустимый порог. В этом случае выбрасывается точка, максимально удаленная (по дуге) от кривой (4), и вся процедура подгонки повторяется. Выброс продолжается до тех пор, пока σ_n не войдет в заданные пределы или пока слишком большое число выброшенных точек не покажет, что кривая вида (3) не может удовлетворительно описать имеющийся набор точек. Обычно такие ситуации возникают, когда в наш набор попадают точки сразу из двух близко расположенных треков, как, напри-

мер, в упоминавшемся выше случае, если оба элемента трека попали в один и тот же интервал гистограммы. Наиболее простой выход из этого положения - проведение повторного гистограммирования с вдвое меньшей шириной полосок. Возможность такого двухэтапного гистограммирования с начальным значением шага по углу $\Delta\theta \sim 1^\circ + 1,5^\circ$, по всем углам для определения областей, где имеются треки, и повторного с шагом $\Delta\theta/2$ в тех местах, где появляются сигналы об опасном сгущении треков, вполне соответствует принципу 2 и существенно экономит общие затраты времени и памяти ЭВМ на гистограммирование (нетрудно подсчитать, что для общего гистограммирования по всем углам с шагом $\theta = 1^\circ$ и вариацией каждого направления в пределах $\pm 3^\circ$ потребуется около 7 тыс. ячеек памяти ЭВМ).

Прослеживание трека с использованием подгонки по кривой вида (3) происходит следующим образом: после того, как по предыдущим точкам были найдены коэффициенты кривой (3) с выбросом всех посторонних точек, находится точка ($R^{(n)}, \theta^{(n)}$) пересечения этой кривой со следующим витком спирали. В окрестности $\pm \delta_{n+1}$ этой точки, ширина которой не превышает $\pm 3 \sigma_n$ (но и не меньше средней ширины трека, т.е. 30 мкм) выбирается точка (R_{n+1}, θ_{n+1}), присоединяемая к массиву точек трека, после чего повторяется вся процедура подгонки кривой и выброса. Если число точек в δ_{n+1} превышает одну, то берется точка, ближайшая к прогнозируемой. Если же в δ_{n+1} точек не находится, то происходит предсказание для следующего витка спирали и т.д., пока величина пропусков не покажет, что трек оборвался или "ушел" в сторону.

Еще более сложная процедура автоматического прослеживания треков на снимках с большой пузырьковой камеры МИРАБЕЛЬ описана в работе /43/.

Использование метода шнуров позволяет объединить этап поиска трековых элементов с этапом прослеживания треков, работая значительно быстрее м.н.к. Однако предсказание по формулам (1-2) надежно работает лишь в благополучных случаях, когда на треке нет больших разрывов и узких пересечений с другими треками.

Можно комбинировать эти два метода, прибегая к экстраполяции по м.н.к. в случаях наличия на треке больших разрывов. Тем не менее возможность ухода на пересекающий под малым углом трек оставалась главным недостатком метода шнуров. В работе /14/ дано обоснование для этапа прослеживания треков с позиций байесовского подхода. Уравнение трека описывается вектором параметров \vec{B} .

На основании информации, полученной по уже найденным точкам трека, делается оценка \vec{B}^* , прогнозирующая положение трека на следующей скан-линии. При этом используется квадратичная функция потерь $I(\vec{B}^*, \vec{B}) = (\vec{B}^* - \vec{B})' (\vec{B}^* - \vec{B})$ и для получения \vec{B}^* минимизируется апостериорный риск. В обычном предположении нормальности вектора измерений в качестве решения получен известный фильтр Калмана. Рассмотрение практической реализуемости фильтра Калмана на ЭВМ потребовало ряда упрощений, сделавших его более быстрым, но суб-оптимальным. Соответствующая программа фильтрации /15/ показала хорошие результаты при обработке реальных событий.

Особую трудность для прослеживания представляют так называемые узкие пары треков, т.е. следов, оставляемых вторичными частицами, выходящими из вершины вместе и расходящимися только в конце поля наблюдения. Для облегчения прослеживания таких узких пар треков оператор должен измерить на концах каждого из них специальную точку (ОТ) так, чтобы обеспечить прослеживание в обратном направлении к вершине.

В случаях отказа обычных методов обнаружения треков может быть применен способ локального преобразования координат с переносом полюса в ОТ. Часть общего массива данных сканирования (R_i, θ_i) в узком секторе, включающем ОТ (R_{OT}, θ_{OT}) , т.е. в интервале значений θ от $\theta_{OT} - \epsilon$ до $\theta_{OT} + \epsilon$, пересчитывается по формулам:

$$\rho = \sqrt{R_{OT}^2 - 2RR_{OT} \cos(\theta - \theta_{OT}) + R^2}$$

$$\varphi = \arcsin\left(\frac{R}{\rho} \sin(\theta - \theta_{OT})\right)$$

в новую систему полярных координат (ρ, φ) с полюсом в ОТ, в которой через полюс проходит только один трек, так что нахождение его упрощенной процедурой проектирования уже не представляет проблемы.

После завершения прослеживания всех треков необходимо отобрать из них только те, которые входят в проекцию события. Этот отбор осуществляется с помощью дискриминантного функционала, составленного с учетом имеющихся априорных сведений. В случае спирального сканирования в качестве проверяемых параметров выбираются:

- расстояние трека от вершины, т.е. параметр β в уравнении (3) и r_1 - полярный радиус ближайшей к вершине точки трека;

- средняя плотность точек вдоль трека, т.е. отношение N/R_N , где N - число точек на треке, R_N - радиус самой далекой от вершины точки.

Если обозначить $P_1 = 1/|\beta|$, $P_2 = 1/R_1$, $P_3 = 1/6$, $P_4 = R_N/N$, то минимизация функционала качества

$$F = - \sum_{i=1}^q W_i P_i \quad (5)$$

на множестве треков-кандидатов позволяет выбрать наилучший из них. Веса W_i подбираются эмпирически так, чтобы члены суммы (5) не слишком отличались друг от друга.

Важным критерием является соответствие кривизны треков, т.е. параметра α в уравнении (3) физическим требованиям, относящимся как к знаку α , так и к величине. Особенно жесткие требования предъявляются к пучковому треку, для которого известны и кривизна и угол ϵ , поэтому оба эти параметра проверяются на нахождение в определенных допусках:

$$|\epsilon_0 - \epsilon_{\text{пучк}}| < \epsilon, \quad |\alpha - \alpha_{\text{пучк}}| < \epsilon_\alpha. \quad (6)$$

Функционал (5) легко обобщается на случай нескольких кандидатов, проходящих критерий (6). Аналогичные способы проверки качества отобранных треков и выбора из них группы, составляющей проекцию события, имеются и для работы с данными в декартовой системе координат.

Дальнейшая обработка направлена на завершение описания события. Точки отобранных треков усредняются, чтобы иметь 15-20 мастер-точек на каждом из них, проверяется комплектность всего события в целом, соответствие всех данных по проекциям и данные переводятся в формат, принятый на входе в программы геометрического и физического анализа.

Допустимые потери во всем этом процессе, называемом обычно фильтрацией, не должны превышать 5-7%, что предъявляет предельно жесткие требования к фильтрации отдельных треков на проекциях. В самом деле, уже для события с пятью треками, измеренными на трех проекциях, требование иметь не более 5% потерь на выходе означает, что каждая проекция трека должна быть обнаружена без ошибки с вероятностью 99,7%.

В этой связи следует подчеркнуть роль проверочных блоков программ фильтрации и, с другой стороны, напомнить, что в некоторых достаточно редких случаях разрешение слишком сложных альтернатив выгоднее предоставить человеку, чем пытаться решить их за счет чрезмерного усложнения программы.

Кроме этого, подключение человека на стадии адаптации программ обработки к новому эксперименту делает систему самообучающейся, так как позволяет немедленно анализировать случаи отказов и ошибок в распознавании и вносить необходимые коррективы в параметры программ. Такое участие человека в оценке результатов фильтрации осуществляется обычно в виде так называемой станции "спасения" событий, представляющей собой средство визуального представления данных в виде дисплея со световым карандашом и соответствующего программного обеспечения, позволяющего "помогать" программе фильтрации и оперативно менять ее параметры^{/10,44/}.

Е. Проблемы обработки изображений при бесфильмовом съеме данных

Несмотря на многие преимущества фильмовой регистрации камерных событий, их главные недостатки - дороговизна и многостепенность обработки - стимулировали поиски других, бесфильмовых методов регистрации, особенно в связи с появлением новых высокостатистических экспериментов с магнитно-искровыми спектрометрами^{/17/} и большими стримерными камерами^{/16/}.

Треки в оптических искровых камерах состоят из коротких прямолинейных светящихся искр (см. рис.8). Стримерные треки при взгляде на них вдоль электростатического поля внешне похожи на треки в пузырьковых камерах (рис.9), но в отличие от пузырьков, имеющих одинаковые диаметры, стримеры имеют разные диаметры и светимость. Выгодное отличие от пузырьковых камер - практически полное отсутствие вторичных взаимодействий, меньший уровень фона и появление полезных событий почти при каждом срабатывании.

Некоторые из современных высокочувствительных телекамер позволяют после необходимого усиления зафиксировать эту светимость искр и стримеров и с помощью соответствующих электронных схем сопряжения передать оцифрованные координаты искр и стримеров прямо в ЭВМ, минуя стадию фотографирования и связанных с ней проблем (освещение, проявка, последующее сканирование и т.д.). Выделяя главные отличия фильмовых и бесфильмовых методов съема информации, укажем прежде всего на принципиальную невозможность при бесфильмовом съеме перемерить событие, которое было измерено или обработано с ошибкой или, что более важно, содержало какой-то редкий эффект и именно поэтому не поддавалось обработке по обычным рутинным алгоритмам.

Будь оно сфотографировано, его можно было бы измерить повторно, исправляя все виды ошибок - как те, что связаны с неправильным выбором шага сканирования и уровня дискриминации, так и ошибки распознавания, прибегая, при необходимости, к участию человека. Хотя подобная надежность не может быть полностью обеспечена при бесфильмовом съеме (т.к. ошибки первого типа возможны и при самом точном контроле отсчетной системы), естественно желание обеспечить максимальное качество воспроизведения изображения путем запоминания всех данных бесфильмового съема на некотором постоянном носителе, например, на магнитной ленте. Однако следует иметь в виду, что само появление бесфильмовых методов было вызвано не столько желанием упростить способ получения и/или удешевить его результат, сколько требованием увеличить статистику и, соответственно, ускорить темп эксперимента. Рост потока данных, доходящего до $10^6 - 10^7$ бит/с требует либо каких-то новых специальных способов видеозаписи с плотностями не менее 300-400 бит/мм (об этой возможности, пока еще малодоступной у нас, но используемой в ЦЕРНе, см. ^{/48/}, см. также ^{/49/}), либо, что более реально, - сжатия данных в процессе их оцифровки и записи на постоянный носитель.

5.1. Сжатие данных оцифровки

Под сжатием данных обычно понимается преобразование, сокращающее их объем так, чтобы оставалась малой некоторая мера разницы между исходным набором данных и реконструкцией сжатого набора. При обработке изображений, состоящих из полезного образа и фоновых данных, эту меру целесообразно строить только на точках, относящихся к полезному образу, что позволяет включить в преобразование сжатия также и процедуру фильтрации (отсева шумов).

Таким образом ставится задача о быстром сжатии без информационных потерь. Это означает, что алгоритм сжатия должен относиться к специальному классу быстрых алгоритмов, допускающих аппаратную реализацию и параллелизм обработки, но сохраняющих при этом с заданной точностью геометрические и физические свойства исходного дискретного представления (более полное изложение этой концепции см. в ^{/45/}). Обычно при этом допускается любое усложнение последующей off-line процедуры реставрации изображения, проводимой по записанным данным, как правило, на мощной ЭВМ. Однако, как излагалось в ^{/46/}, если и алгоритм реставрации исходного изображения также удастся реализовать как быстрый, то возникает возможность адаптивного сжатия, т.е. такой динамической перестройки параметров сжатия, которая его улучшает, уменьшая меру разности исходного и восстановленного изображений (или их фрагментов).

Что же касается требования сжатия без информационных потерь, то это означает не только условие сохранения всей информации, необходимой для надежного прослеживания треков, но также и дополнительное условие, специфичное, например, для стримерных треков при высоких энергиях, а именно - сохранение информации об их локальной структуре, то есть о соотношении отсчетов и пропусков вдоль каждого трека, позволяющем оценить их первичную ионизацию (см. результаты экспериментальных и теоретических исследований в ^{/23-25/}).

В такой постановке задача быстрого сжатия исключает возможность простого перенесения того богатства алгоритмов распознавания треков, которое было накоплено в пленочной обработке, поскольку сам подход при распознавании означал выделение множества точек, относящихся к полезному образу, и отбрасывание остальных как фоновых. Заметим, что здесь мы не рассматриваем проблемы сжатия для бесфильмового съема данных с искровых спектрометров, поскольку потоки данных там (с видиконной системы спектрометра "ОМЕГА", например ^{/17/}) позволяют фиксировать все на обычную магнитную ленту. Это дает возможность применять затем в режиме off-line параметрический метод распознавания, описанный в начале нашей лекции как метод главных компонент. Для преодоления отмеченных выше трудностей, связанных главным образом с пропусками на треках и неоднородностью магнитного поля, был предложен ^{/37/} оригинальный метод двумерного поиска треков. Треки в боковых камерах, окружающих мишень (см. рис.8), использовались для вычисления пространственного положения вершины путем продолжения их до пересечения с пучком, направление которого регистрировалось специальными проволочными камерами. Прослеживание треков в передних камерах велось от наиболее удаленного от мишени конца назад к ней. Прослеживание велось обычным методом подгонки по м.н.к. траектории по пяти последовательным искровым промежуткам (в трех из которых по крайней мере должны присутствовать искры) с предсказанием положения искры в последующем искровом промежутке. Для сшивания треков осуществлялся выход в пространство (трек-матч) оригинальным методом, предложенным в ^{/22/}.

Подобную ситуацию мы имеем и для стримерных треков в случае, когда информация об их локальной структуре не является необходимой, и достаточно сохранять информацию, требуемую только для геометрической программы. Становится возможным использовать методы сжатия путем быстрого отслеживания треков или применения быстрых ортогональных преобразований. Варианты такого быстрого сжатия для стримерной камеры установки "РИСК", описанные в ^{/18/}, позволяют с помощью аппаратно реализованных блоков слежения по методу "шнуров" добиться высокого коэффициента сжатия (до 20+30) за счет выполняемого в процессе

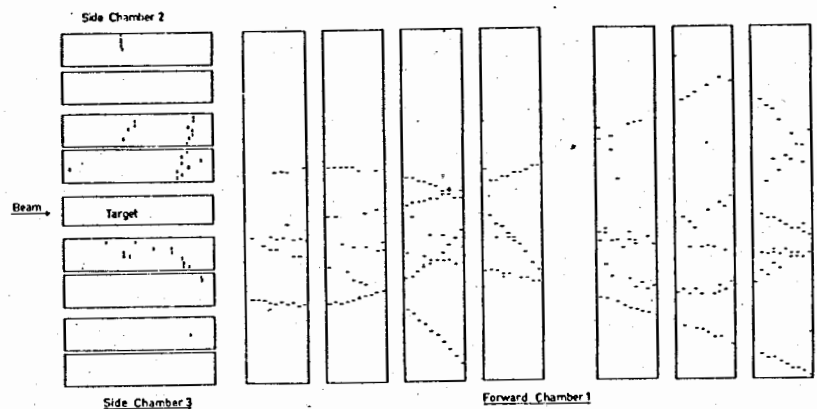


Рис.8. ^{/17/}.

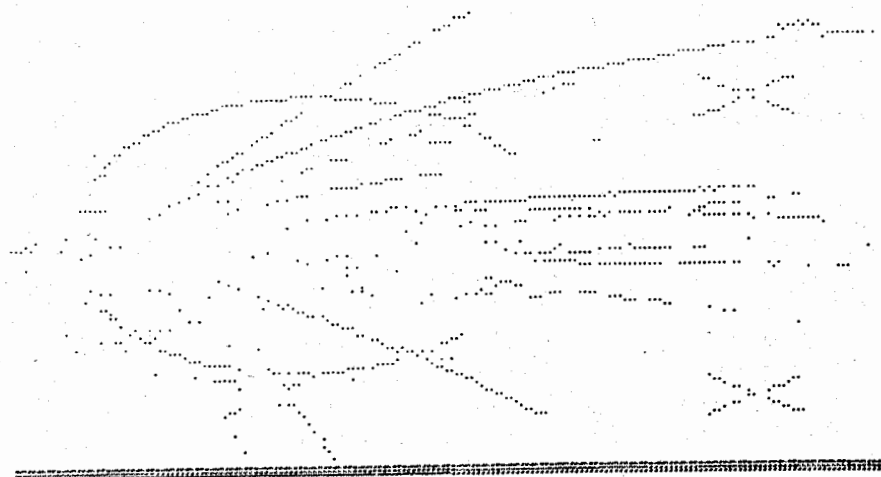


Рис.9.

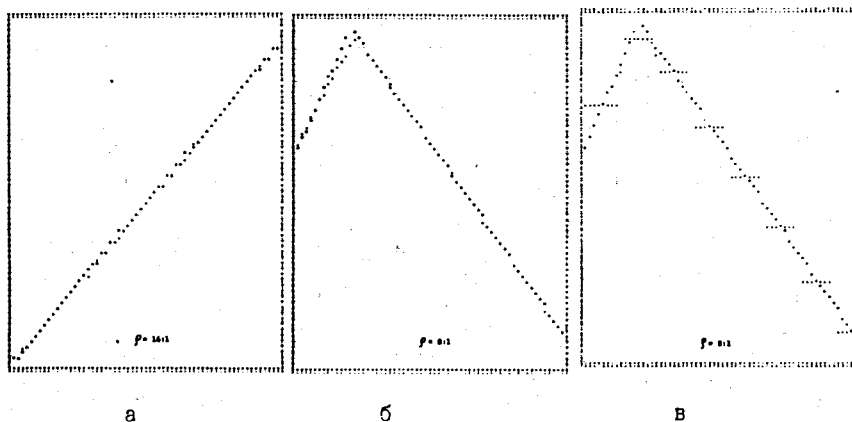


Рис. 10.

слежения заполнения промежутков на треках и отсева шумов. Если те же операции фильтрации и пополнения пропусков выполнять в виде специальной быстрой процедуры предварительной сортировки данных, то возникает возможность использовать для сжатия быстрые ортогональные преобразования с последующим урезанием спектра, как это делалось в методе главных компонент. Но в отличие от используемого там разложения Карунен-Лозва, требующего больших предварительных вычислений для определения корреляционной матрицы, было найдено^{47/}, что для двумерных изображений более удобным является быстрое слэнт-преобразование (БСП)^{48/}.

В состав базовых функций БСП входит наклонная прямая, что позволяет ограничиться всего 8+16 первыми гармониками спектра для представления кусочно-линейных фрагментов двумерных изображений. На рис. 10 представлены данные применения БСП и результат реконструкции прямой и ломаной линий. Через ρ обозначен коэффициент сжатия. Для сравнения на рис. 10в приведены данные по применению к тому же изображению известного преобразования Уолша-Адамара, подтверждающие оптимальность БСП.

Однако в общем случае быстрое сжатие может быть осуществлено только путем объединения в сегменты подряд идущих отсчетов на прямолинейных участках треков, причем случаи заполнения небольших промежутков должны обязательно сопровождаться запоминанием сведений о них, равно как и о выброшенных шумовых точках, что должно давать a priori меньший коэффициент сжатия.

Как и в случае спецпроцессоров типа SNR^{29/}, рассчитанных на сжатие пленочной информации в процессе сканирования снимков, требование быстродействия алгоритмов ведет к необходимости разбиения изображения на фрагменты – слайсы, объединяющие координаты, полученные на n последовательных скан-линиях, в пределах которых и производится быстрая проверка связности точек. Если число точек n в такой связанной группе превышает некоторый небольшой порог, то в пределах слайса эти точки аппроксимируются линейным трек-элементом (ЛТЭ). Сжатие данных обеспечивается тем, что вместо указания координат всех точек группы выдаются только 3+5 параметров этого ЛТЭ. Очевидное искушение: увеличить длину слайса N , чтобы поднять коэффициент сжатия, – очень опасно, т.к. при этом из-за кривизны трека неизбежно будет нарушена заданная точность δ линейной аппроксимации. Поэтому в спецпроцессорах такого рода N выбирается достаточно малым (5+10).

Однако в случае бесфильмового съема значительное увеличение длины слайсов оказалось необходимым, чтобы компенсировать a priori большее число параметров ЛТЭ. Линейная аппроксимация с заданной точностью δ при этом естественным образом стала кусочно-линейной, т.е. для представления некоторой отслезенной связанной группы точек, соответствующих какому-то фрагменту трека, могло потребоваться более, чем один, ЛТЭ, что опять-таки вело к снижению коэффициента сжатия (не говоря уже о требовании сохранения координат всех неидентифицированных точек).

Таким образом, для разработки конкретных алгоритмов быстрого сжатия без информационных потерь в условиях бесфильмового съема было необходимо разрешить противоречия между требованиями к получению достаточно высокого коэффициента сжатия и точности аппроксимации при заданном малом времени (~ 65 мс/кадр) работы всего алгоритма после его реализации в виде аппаратуры сжатия данных (АСД). С этой целью процедура сжатия была разбита на два этапа: 1) отслеживание связанных точек путем быстрой сортировки данных сканирования в пределах каждой скан-линии, и 2) кусочно-линейной аппроксимации с заданной величиной максимального отклонения δ в комбинации с запоминанием координат неидентифицированных точек (подробнее описание программ, имитирующих работу АСД, дано в^{50/}).

Кроме этого, для выбора оптимальной длины слайсов и других параметров АСД было необходимо выработать критерии качества сжатия и провести исследование на реальных данных бесфильмового съема. Такой материал был получен с помощью записи на магнитную ленту достаточно представительного набора оцифрованных изображений с телевизионной системы контроля^{53/} стримерной камеры установки "РИСК".

Как видно из рис.9, представляющего одно из таких оцифрованных изображений, число точек на нем достаточно мало (1+2 тыс.), что объясняется уже упоминавшимся отсутствием треков вторичных и непровзаимодействовавших первичных частиц, а также в области мишени. Кроме того, на большинстве проекций, не содержащих изображения области взаимодействия, кривизна треков мала, что должно обеспечивать высокую точность их линейной аппроксимации на участках значительной длины (вплоть до 64 скан-линий).

В качестве оптимизирующих критериев качества сжатия были предложены следующие характеристики: коэффициент сжатия, среднеквадратичная величина отклонения точек исходного изображения от соответствующих им ЛТЭ, а также средние на кадр относительные числа неидентифицированных точек и всех полученных ЛТЭ (оба по отношению к числу точек на исходном оцифрованном изображении).

Исследование^{51/}, проведенное на указанных реальных данных, показало, что при точности восстановления треков 2,5 мм (в пространстве камеры) АСД позволяет добиться 8-кратного сокращения исходных данных, а при сохранении координат неидентифицированных точек - 3-кратного. При этом использовались длины слайсов в 32+64 скан-линии (при 512 скан-линиях на кадр). Некоторое улучшение показателей сжатия было достигнуто затем путем такого усовершенствования алгоритма специальной сортировки, которое обеспечивает учет локального направления фрагмента трека по уже найденным его точкам. Оценки результатов сжатия тех же реальных данных с помощью такого мощного и универсального спецпроцессора, как SHP, показали, что более скромный по возможностям, но специально ориентированный на бесфильмовый съем, АСД не уступает SHP с точки зрения введенных критериев сжатия^{52/}.

5.2. Обработка сжатой информации

Возвращаясь к основной проблеме обработки изображений, напомним, что в рамках общей схемы нахождения и описания событий выделение линейных треков-элементов, выполняемое в процессе сжатия данных, предваряет следующие два этапа: распознавание проекций треков и объединение их в события. Как это видно из рис.11, где изображен набор ЛТЭ, полученных при сжатии данных, представленных на рис.9, используемые для этой цели методы должны быть отнесены к классу методов без наведения, причем в силу специфики стримерной камеры и бесфильмовой системы регистрации возникает ряд дополнительных трудностей:

- принципиальное отсутствие данных о вершине события и прилегающей к ней области, занятой мишенью;
- значительные оптические искажения, вносимые электронно-оптическими преобразователями (это препятствует использованию обычной

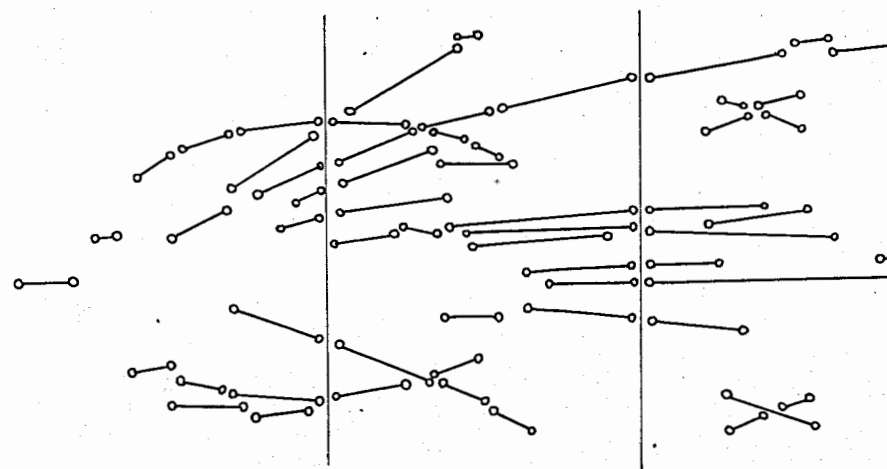


Рис.11.

методики применения уравнения окружности для описания проекции трека);

- необходимость использования массива неидентифицированных точек для поиска треков, наклон которых превышает 45° ;

- отсутствие соответствия между порядком следования ЛТЭ внутри слайса и треками события и т.д.

Тем не менее сокращение данных до 1+2 сот ЛТЭ/кадр значительно повышает их информативность и делает вполне реальным применение на ЭВМ среднего класса глобальных методов распознавания, требующих введения на множестве всех возможных ЛТЭ некоторой метрики, т.е. расстояния, характеризующего близость ЛТЭ с точки зрения их принадлежности к некоторой неизвестной пока выпуклой кривой, являющейся моделью проекции трека.

При определении такого расстояния необходимо учитывать как возможные малые ошибки в определении концов ЛТЭ, так и геометрические факторы, определяющие эту близость. С этой целью рассмотрим возможные случаи взаимного расположения двух ЛТЭ на плоскости и установим некоторые очевидные критерии их близости.

1. Близкие отрезки не должны перекрываться или, наоборот, быть слишком удалены друг от друга, т.е. должен быть справедлив критерий соответствия их внутренних концов (точек b и c на рис.12), который количественно может быть выражен как требование к величине $x = x_3 - x_2$:

$$0 < \Delta x < x_{LIM}. \quad (7)$$

2. В случае невыпуклой конфигурации двух отрезков (рис.12а) их близость означает, что они коллинеарны с некоторой допустимой неточностью, т.е.

$$\Delta y_{max} = \max(\Delta y_1, \Delta y_2) < \delta. \quad (8)$$

3. При выпуклой (вогнутой) конфигурации (рис.12б), когда $\Delta y_1 \cdot \Delta y_2 > 0$, близость отрезков означает их близость к проекции трека, приближенной моделью которой в случае достаточно малого x_{LIM} может служить дуга окружности. Чтобы проверить близость двух отрезков к одной и той же окружности, достаточно найти точку O пересечения перпендикуляров к серединам отрезков и убедиться в примерном равенстве радиусов R_1 и R_2 , т.е. расстояний от точки O до концов отрезков. Проверяемое условие:

$$\Delta R = |R_1 - R_2| < \Delta R_{LIM}. \quad (9)$$

Чтобы предотвратить поиски продолжения изолированных ЛТЭ, являющихся, например, фиктивными, т.е. порожденными случайным скоплением шумовых точек, необходимо ввести правило запрета на те случаи, когда условия (7)-(9) оказываются нарушенными.

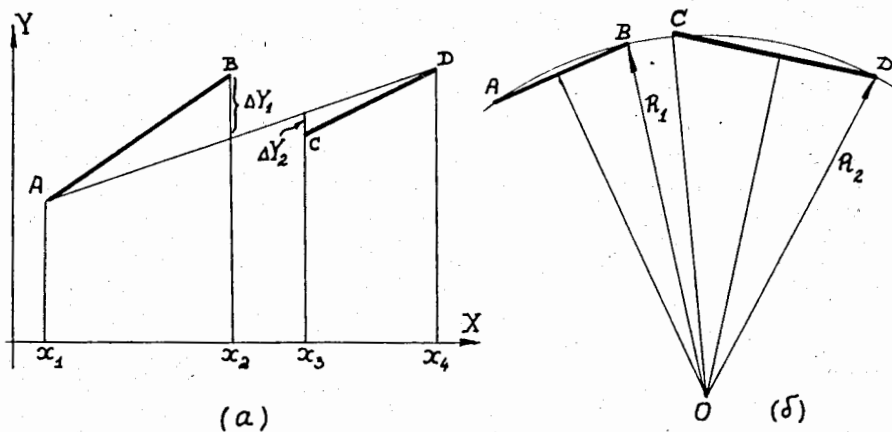


Рис. 12

Для объединения всех этих критериев и правила запрета в одну формулу, определяющую меру ρ близости двух ЛТЭ, воспользуемся функцией Хевисайда

$$\chi(x, a) = \begin{cases} 1, & x < a \\ 0, & x \geq a \end{cases} \quad (10)$$

После этого на основании (7)+(10) получаем

$$\rho = w_1 \cdot \Delta x \cdot \chi(\Delta x, x_{LIM}) \cdot \chi(-\Delta x, 0) + w_2 \cdot \Delta y \cdot \chi(\Delta y, \delta) + w_3 \cdot \Delta R \cdot \chi(\Delta R, \Delta R_{LIM}) \cdot (1 - \chi(\Delta y, \delta)) + m [1 - \chi(\Delta R, \Delta R_{LIM}) + (1 - \chi(\Delta x, x_{LIM})) \cdot (1 - \chi(-\Delta x, 0)) + (1 - \chi(\Delta y, \delta))] \quad (11)$$

Здесь веса w_1, w_2, w_3 , в сумме равные единице, а также параметры x_{LIM}, δ, R_{LIM} подбираются по обучающей выборке изображений. Множитель m перед последним членом должен быть выбран достаточно большим ($> 10^4$), чтобы обеспечить выполнение правила запрета.

Используя функцию ρ для проверки взаимной близости всех ЛТЭ, можно построить симметричную матрицу близости и известными методами /54,55/ выделить наборы ЛТЭ, соответствующие отдельным трекам. Пример удачного применения такого глобального метода для обработки слабых фильмовых данных на ЭВМ БЭСМ-6 дан в работе /56/.

Однако, если учесть вышеуказанные отличия ЛТЭ в бесфильмовом случае и принять во внимание идеи, наиболее явно высказанные в статье Д.В. Матулы в книге /38/ "...Представление исходных данных о близости в виде графа близости, а не в виде симметрической матрицы, построенной из действительных чисел, явно выгодно с вычислительной точки зрения", то по-видимому, более рациональными для расчетов на ЭВМ можно считать методы кластеризации на графах близости.

Рассмотрим плоский взвешенный граф, вершины которого образованы концами всех ЛТЭ. Ребра, совпадающие с ЛТЭ, будем считать заданными, т.е. имеющими нулевой вес (длину). Задавая веса остальных ребер с помощью функции ρ , будем искать минимальное связующее дерево (MCD) этого графа, используя известный алгоритм Прима /57/. Появление в последние годы программных реализаций этого алгоритма, требующих на порядок меньшее число операций и не нуждающихся в рабочей памяти ЭВМ (см., например, /58/), - еще один довод в пользу глобальных методов, основанных на теории графов.

Правило запрета обеспечивает разделение MCD на несколько ветвей -кластеров путем устранения несущественных (имеющих слишком большой вес) ребер. Полная процедура разбиения MCD на ветви, соответствующие отдельным трекам, производится обычным методом кластеризации по средним значениям веса кластеров с отбрасыванием ребер, выходящих по своему весу за три среднеквадратичные ошибки. Поскольку детали этого метода описаны (см. напр. /59/), отметим только, что по сравнению с

обработкой "сырых", несжатых данных сканирования (см./13/) учет направления ЛТЭ в весовой функции ρ значительно усилил эффективность этого метода при разрешении трудностей с треками, пересекающимися под малыми углами, или близко идущих параллельными треками.

Одним из последствий универсальности глобальных методов распознавания являются значительные затраты ресурсов ЭВМ и, прежде всего, памяти на их реализацию. Поэтому так же, как при обработке несжатой информации, более экономными в реализации оказываются методы прослеживания, позволяющие, кроме всего прочего, сократить область перебора кандидатов за счет лучшего учета специфики обрабатываемых данных.

Прослеживание или односвязывающий метод с поиском ближайшего соседа^{38/} (в терминах кластерного анализа) – это процедура, в которой, начиная от выбранного исходного элемента, осуществляется поиск конца линии путем присоединения ближайшего (с точки зрения функции расстояния ρ) элемента.

Естественные требования к такой процедуре: прослеживание должно быть достаточно "безразличным" к случайным отклонениям элементов "своего" кластера, чтобы доходить до конца линии, но в то же время – достаточно "четким", чтобы распознавать изолированные шумовые элементы или элементы, относящиеся к другим линиям, и не присоединять их к кластеру, что неизбежно ведет либо к обрыву прослеживания, либо к "уходу" на другую линию.

Противоречивость этих требований может быть разрешена путем уточнения понятия близости с ростом кластера. Для этого функция ρ должна быть дополнена еще и операцией точного прогнозирования того района, в котором следует искать следующий элемент кластера с помощью экстраполяции по уже найденным элементам. Следует подчеркнуть существенное отличие в постановке задачи экстраполяции по набору ЛТЭ от экстраполяции по точкам. Экстраполирующая кривая (обычно это – парабола или окружность) должна быть не просто проведена по точкам, определяющим концы ЛТЭ (или через их середины с сохранением близости производной кривой к направлениям ЛТЭ). Должно выполняться требование, положенное в основу идеи о замене изображения трека набором ЛТЭ, а именно: фитирующая кривая не должна выходить за границы коридоров, окружающих каждый ЛТЭ в соответствии с процедурой кусочно-линейной аппроксимации с заданной точностью.

По-видимому, не существует аналитического решения этой задачи для окружности в такой именно постановке, исключающей любой выход окружности за пределы полос фиксированной ширины. Во всяком случае уже для трех ЛТЭ можно построить пример вполне обычной ситуации, когда решения может не существовать из-за случайного разброса концов

ЛТЭ. Однако, если перейти к более реалистическому условию попадания кривой в набор полос переменной ширины с заданной доверительной вероятностью $1 - \alpha$ (α – мало), то задача определения коэффициентов полинома, наименее уклоняющегося от всех ЛТЭ в метрике L_2 , может быть решена методом неопределенных множителей Лагранжа. Аналитическое решение достаточно громоздко и, кроме того, использование для экстраполяции полинома уже второй степени за пределами того участка, где он задан, должно проводиться с большой осторожностью.

В этом отношении экстраполяция на небольших расстояниях с помощью кривой постоянной кривизны – окружности – явно предпочтительнее. На практике хорошо себе зарекомендовала следующая простая вычислительная процедура нахождения экстраполирующей окружности по трем ЛТЭ (на большей длине окружность перестает быть адекватной моделью трека). Вначале делается перевод шести точек, являющихся концами ЛТЭ в локальную систему координат (x^*, y^*) с центром в центре их тяжести и ось абсцисс, проведенной вдоль этих точек по методу наименьших квадратов (м.н.к.). В новой системе координат строится парабола м.н.к. вида $y^* = ax^{*2} + b$. Радиус окружности определяется, как $R = 1/2a$, а координата центра y_c^* также подбирается по м.н.к. Вычисляются отклонения от этой окружности в концах и серединах ЛТЭ и находится максимальное из них – D_{max} . Если оно оказывается меньше, чем заданная величина δ , то процедура оканчивается, и координаты центра $x_c^* = 0$ и y_c^* переводятся в исходную систему координат. При $\delta < D_{max} < 3\delta$ делается соответствующая коррекция радиуса и повторно вычисляется D_{max} . Если оно не убывает, выдается диагностика о невозможности проведения окружности.

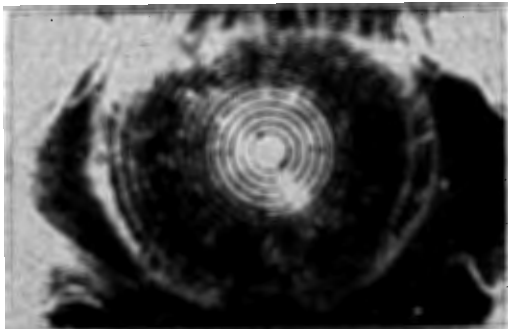
После нахождения в предсказанном районе очередного ЛТЭ его близость к последним двум из уже найденных также проверяется по этой процедуре.

Аналогичная проблема экстраполяции возникает и на следующем этапе распознавания: объединения треков в событие и отсева фоновых и фиктивных треков. Экстраполируя вторичные треки назад до пересечения с продолжением пучкового трека, радиус которого известен, можно оценить координаты вершины как центра тяжести точек пересечения, взвешенных с учетом некоторой функции качества треков, включающей их длину, заполненность отсчетными точками и т.д. Фрагменты треков, продолжение которых не проходит вблизи вершины, метятся в банке треков как фоновые.

Определение примерных координат вершины необходимо также для проведения поиска среди неидентифицированных точек тех возможных треков, которые не были обнаружены из-за слишком большого угла (свыше

Рис.13.

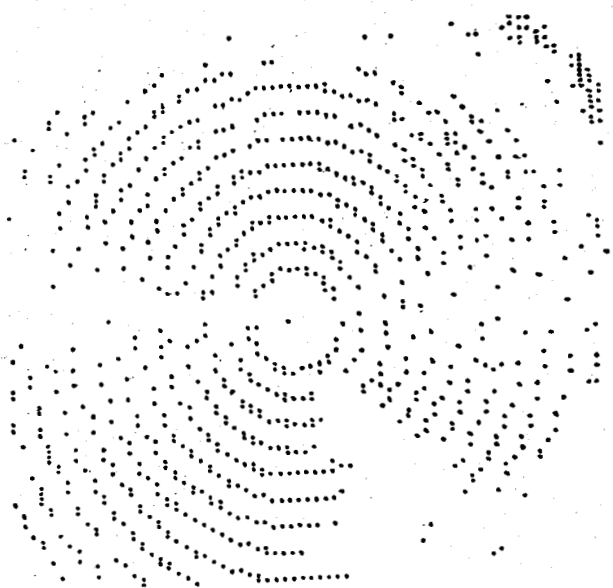
- а) сферическая роговица
- б) деформированная роговица
- в) результаты оцифровки роговицы (а)



а)



б)



в)

45°) с направлением пучкового трека. Кроме того, должны быть сделаны попытки использования неидентифицированных точек для продолжения к вершине и от нее уже найденных коротких сегментов треков и, возможно, присоединения дополнительных точек к остальным трекам.

Окончательное формирование банка всех треков события может быть выполнено только после "сшивания" и сопоставления всех проекций с выходом в пространство, что делается так же, как и в случае фильмовой обработки и по тем же самым программам^{/60/}.

В заключение этого раздела подчеркнем преимущества введения дополнительного уровня распознавания изображений – выделения линейных трек-элементов. Появившийся в силу необходимости сжатия данных этот уровень оказался затем вполне оправданным с точки зрения повышения эффективности и качества процедур распознавания на последующих уровнях.

Одним из свидетельств тому может служить, в частности, работа^{/61/} о методах обработки данных сканирования фотографий глаза человека с нанесенными на них оптическим путем кольцевых реперных меток (см. фото на рис.13). Целью автоматической обработки являлось нахождение двух взаимно-перпендикулярных главных диаметров этих колец и вычисление с высокой точностью (~6 мкм) координат точек пересечения колец с этими диаметрами.

После перевода данных оцифровки изображения в полярную систему координат кольца перешли в набор волнообразных линий с характерными разрывами (см. рис.14). Каждая из этих линий была отслежена и заменена на 12 линейных элементов, по которым были проведены кривые вида $R=R_0 + A\cos 2\varphi + B\sin 2\varphi$, позволившие проинтерполировать недостающие участки, вычислить положение главных диаметров и после возврата в исходную декартову систему координат получить координаты искомым точек с требуемой точностью уже при 256 скан-линиях/кадр.

6. Заключение

Проблемы цифровой обработки изображений, возникающие в физике, успешно решаются в Объединенном институте ядерных исследований и других физических центрах, где работают устройства для автоматической обработки снимков с пузырьковых и искровых камер на базе НРД и СИ, эксплуатируются системы для бесфильмового съема данных с искровых и проволочных камер и ведутся исследования по использованию телевизионной аппаратуры для бесфильмового съема данных. Эти, а также и не затронутые здесь работы по электронно-лучевому сканированию освещены в трудах международных симпозиумов и конференций^{/26,27,62/}.

Развитие новой экспериментальной техники в физике высоких энергий ставит новые задачи перед специалистами по цифровой обработке изображений. Перечислим наиболее актуальные из них.

Повышение энергии пучков частиц до сотен и тысяч млрд электрон-вольт ведет к резкому возрастанию числа вторичных частиц, идущих вперед в узком конусе, среди которых растет доля нейтральных частиц. До определенных технических и экономических пределов возникающие проблемы могли быть решены на традиционных путях применения методики пузырьковых камер за счет увеличения их размеров и использования широкоформатной оптики. Это привело к созданию семейства гигантских пузырьковых камер и соответствующих программных комплексов для обработки получаемых стереофотографий, либо на "старых" автоматах типа HPD или SR после специальной модернизации аппаратуры и перестройки программ распознавания с обязательной ориентацией на участие человека на стадии "спасения" событий, не прошедших обработку, либо на автоматизированных измерителях нового типа, основанных, как правило, на применении сканирующих ЭЛТ, особенно подходящих для включения человека в сам процесс распознавания.

Однако растущий интерес к изучению новых физических свойств короткоживущих элементарных частиц требовал новой экспериментальной техники, которая позволяла бы увеличить число срабатываний за цикл ускорителя, заполнять пространство детектора значительно большим числом частиц и в то же время сводить к минимуму фоновые процессы. В этой связи быстро возросла роль электронных экспериментов, больших искровых спектрометров, стримерных камер и целых комплексов из них - гибридных систем, состоящих из быстроциклирующих пузырьковых камер, используемых в качестве мишени-детектора, и окружающих их по секторам наиболее вероятного выхода продуктов взаимодействия систем из стримерных и искровых камер. Новинка последних лет - голографические камеры, - обеспечивают чрезвычайно высокое пространственное разрешение (~ 10 мкм) при очень большой глубине резкости, что позволяет заполнить камеру огромным числом частиц, резко повышая ее эффективность.

Два главных фактора: 1) резкое возрастание потоков обрабатываемой информации и 2) сложность распознавания, делают невозможным организацию всей обработки только традиционными средствами. Уже на примере блоков распознавания и фильтрации программных комплексов, предназначенных для автоматической обработки снимков с гигантских пузырьковых камер, можно было проследить основные тенденции развития этих программ:

- усложнение моделей треков;
- увеличение числа пороговых процедур, требуемых для решения многочисленных вариантов прослеживания при "перепутывании" треков;

- необходимость в специальной процедуре "сшивания" (bridging) фрагментов треков, наблюдаемых "по частям" разными объективами;

- увеличение уровней распознавания за счет как уже упоминавшихся способов префильтрации, так и, главным образом, специальных программных или аппаратных средств, заменяющих множество точек изображения на значительно меньшее число линейных трек-элементов;

- резкое возрастание ориентации на помощь человека в запутанных случаях распознавания.

О необходимости применения подобных приемов развития методов распознавания данных больших стримерных камер и искровых спектрометров свидетельствовали материалы предыдущего раздела данной лекции и другие работы^{/60,63/}

Имеются оценки^{/28/}, показывающие возможность применения автомата SR для обработки информации от быстроциклирующей пузырьковой камеры, входящей в гибридную систему EHS, при энергиях свыше 200 ГэВ/с, но при условии получения целеуказаний от других устройств, обрабатывающих остальную часть информации.

Дальнейший обзор приемов распознавания, достаточно общих для автоматической обработки данных в физике высоких энергий, затруднен как спецификой, накладываемой новейшей экспериментальной техникой, так и многообразием способов отбора и переработки получаемой информации, определяемых этой спецификой и последними достижениями вычислительной техники, включая средства представления данных. При этом меняется сам подход к распознаванию. Оно, как кажется, перестает быть автоматическим, т.к. в нем участвует человек, но это участие на совершенно новом уровне, позволяющем в необходимых случаях максимально использовать распознавательные способности человека, ускоряя при этом в сотни и тысячи раз исполнение остальных функций, второстепенных с точки зрения затрат времени на распознавание и усилий на программирование. Таким образом, общие направления усилий по решению проблем, вызываемых факторами (1)-(2), могут быть сведены к следующим:

1. Возрастание роли специальных аппаратных средств быстрой селекции и обработки в целях сжатия данных без потери физической информации. Основное направление - быстрое выделение элементов треков как линейных^{/29-31,52/}, так и более сложной структуры^{/64/}.

2. Возрастание роли участия человека в процессе распознавания как на заключительном этапе обработки с помощью упоминавшихся в § 4 станций спасения, так и на качественно новом уровне, когда человеку даются средства для более наглядного и пространственного представления и активного вмешательства в процесс распознавания с помощью включения программы геометрической реконструкции в число программ, рабо-

таких на линии с оператором, и усовершенствования программ и средств дисплейной техники.

3. Учет появления новых средств ввода изображения в ЭВМ. Кроме устройств телевизионного съема информации, следует указать на еще более перспективные приборы с зарядовой связью (ПЗС), которые позволяют высокоточную оцифровку не только бинарной^{/65/}, но и полутоновой информации (см. напр.^{/66/}), и могут быть, в частности, использованы для определения ионизации стримерных треков^{/32/}.

4. Использование последних достижений вычислительной техники от микропроцессоров до сверхбыстрых ЭВМ параллельного действия. Появление дешевых одно - двухбайтовых микропроцессоров с развитой системой команд и других схем большой интеграции дало новые возможности для перевода в режим он-лайн многих операций контроля и селекции данных (чем во многом объясняется рост удельного веса электронных экспериментов), а функциональная гибкость мультисистем из микро-ЭВМ позволила применять их для быстрой параллельной обработки бинарных и полутоновых изображений^{/70/}.

Как мы помним, целесообразность применения параллельных вычислений на всех уровнях распознавания изображений доказывалась первыми опытами на уникальной ЭВМ ИЛЛИАК-III около двух десятилетий назад. Однако только теперь эта идея получает качественно новую техническую базу в виде таких средств ввода изображений, как матрицы ПЗС и сравнительно дешевых процессоров параллельного действия (так называемых вычислительных сред^{/33/}, векторных и ассоциативных процессоров^{/67/}), включение которых в состав обычных малых ЭВМ многократно увеличивает их скорость и вычислительные мощности, но для определенного круга задач, в том числе и задач распознавания образов. Большие перспективы возникают для проблем создания параллельных алгоритмов в задачах, появляющихся на высших уровнях распознавания (таких, например, как пространственная реконструкция треков во взаимодействиях, где рождаются несколько десятков частиц). Разработка этих алгоритмов необходима для использования специализированных параллельных процессоров, а также и для решения таких задач на универсальных сверхбыстрых ЭВМ параллельного действия (типа КРЭИ-I или CDC-205^{/68/}).

Привлечение огромных государственных средств для развития фундаментальных наук и в первую очередь физики позволило сосредоточить в этой области большие коллективы первоклассных специалистов и снабдить их самой новейшей техникой. Одним из последствий этого явилось опережающее по сравнению с другими областями науки и техники развитие средств электроники, вычислительной техники и математических методов обработки экспериментальных данных. Особенно отчетливо это видно на

примере приведенных в настоящей лекции достижений в области распознавания образов в физике высоких энергий, что, в частности, объясняется сравнительно простым характером изображений. Немалую роль здесь сыграл и международный характер исследований.

В этой связи вполне естественными являются идеи о применении накопленного в экспериментальной физике богатства алгоритмов и техники в других областях науки и народного хозяйства: в медицине, биологии, металлургии, для обработки космических наблюдений и т.д. В наших лекциях мы ограничимся примером с обработкой медицинской информации, приведенным в конце предыдущего раздела. Большое количество других интересных применений можно найти в трудах семинара^{/26/} и конференций^{/62,69/}.

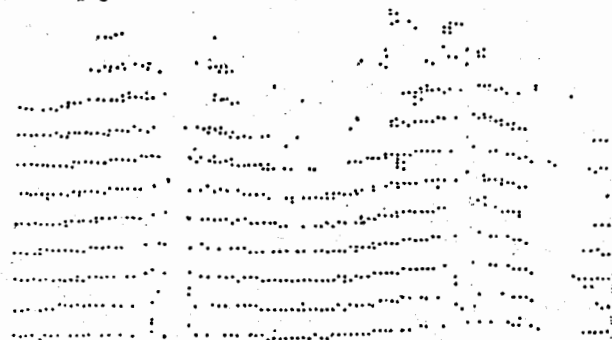


Рис. 14

Литература

1. Ососков Г.А. ОИЯИ, 5-9787, Дубна, 1976.
2. Фу К. Структурные методы в распознавании образов. "Мир", М., 1977.
3. Нарасимхан Р. В сб.: Автоматический анализ сложных изображений. "Мир", М., 1969, с. 22.
4. Автоматическая обработка данных с пузырьковых и искровых камер. (Под редакцией Б.С.Розова), Атомиздат, М., 1971.
5. Ососков Г.Н. ОИЯИ, П10-83-187, Дубна, 1983.
6. Стрэнд Р. В сб.: Распознавание образов при помощи ЦВМ. "Мир", М., 1974.
7. Международный симпозиум по вопросам автоматизации обработки данных с пузырьковых и искровых камер. ОИЯИ, Д10-6142, Дубна, 1971.
8. H.White, IEEE Trans. Comput., C-20, No.9, 1971.
9. L.Zanello, CERN 72-16, 1972, p. 141.
10. Нергуй Б., Ососков Г.А. В кн.: Совещание по программированию и математическим методам решения физических задач. ОИЯИ, Д10-II-II264, Дубна, 1978, с. 460.
11. Байла И., Ососков Г.А. ОИЯИ, П10-II834, Дубна, 1978.
12. Ососков Г.А. Материалы семинара по обработке физической информации. Изд. ЕРФИ, Ереван, 1976, с. 268.

13. Ососков Г.А., Пахомов В.Л. В кн.: Совещание по программированию и математическим методам решения физических задач. ОИЯИ, Д10-II-11264, Дубна, 1978, с. 288.
14. Локтионов А.А. Труды ИФВЭ АН КазССР, т.2, Алма-Ата, 1974, с. 138.
15. Локтионов А.А. и др. Известия АН КазССР, серия физ.-мат. № 2, Алма-Ата, 1978.
16. G.Bohm et. al., Proc. of the 1st International Conf. on Streamer Chamber Technologs, ANL, 1972, p. 117.
17. H.Grote et. al., CERN dd/72/73.
18. Ососков Г.А. и др. ОИЯИ, P10-II213, Дубна, 1978.
19. Алексеева Н.П. и др. В кн.: Материалы Всесоюзного семинара по обработке физической информации. Изд.ЕрФИ, Ереван, 1976, с.420-425.
20. Автоматический анализ сложных изображений. (Сб. под редакцией Э.М.Браввермана), "Мир", М., 1969.
21. Хуанг Т. и др. В сб.: Обработка изображений при помощи ЭВМ. "Мир", М., 1973, с. 9.
22. G.E. Bourgeois, Comp. ACM 14, 1971, p. 802.
23. Кадыков Г.М. и др. ОИЯИ, Д10-7707, Дубна, 1974.
24. Кулюкина Л.А. и др. ОИЯИ, P5-III43, Дубна, 1977.
25. Гласнек К.П., Кадыков Г.М. ОИЯИ, 10-10923, Дубна, 1977.
26. Совещание по программированию и математическим методам решения физических задач. ОИЯИ, Д10, II-11264, Дубна, 1978.
27. Труды II семинара по обработке физической информации. Изд.ЕрФИ, Ереван, 1978.
28. J.Garavel et. al., CERN/EP/ENS/1978.
29. Mazetti M., Nucl. Instr. Meth., 135, 1976, pp.427-433.
30. W.Jank, CERN 74-23, 1974, pp.367-401.
31. Федотов О.П. В сб.: Материалы семинара по обработке физической информации. Изд.ЕрФИ, Ереван, 1976, с.198.
32. Ваттенбах К. и др. В сб.: Тезисы II Всесоюзного совещания по автоматизации научных исследований в ядерной физике. "Наука" КазССР, Алма-Ата, 1978, с. 57.
33. The FPS AP-120B Array Processing, Floating Point Syst. Inc. (фирменные материалы).
34. Говорун Н.Н., Дикусар Н.Д. В кн.: Международный симпозиум по вопросам автоматизации и обработки данных с пузырьковых и искровых камер. ОИЯИ, Д10-6142, Дубна, 1971, с.525.
35. Головин Б.Н., Кулюкина Л.А. ОИЯИ, 2348, Дубна, 1965.
36. Wind H. CERN 72-21, 1972, pp. 53-106.
37. Hansroul M. et.al. В кн.: Совещание по методам программирования и решению физических задач. ОИЯИ, Д10-7707, Дубна, 1974, с.460.

38. Классификация и кластер (сб. под редакцией Дж.Вэн.Райзена), "Мир", М., 1980.
39. Котов В.М. и др. ОИЯИ, 10-7939, Дубна, 1974.
40. Р.Позе. В кн.: Материалы семинара по обработке физической информации. Изд.ЕрФИ, Ереван, 1976, с.326.
41. Абдурахимов А.У. и др. ОИЯИ, 10-80-831, Дубна, 1980.
42. Нергуй В. ОИЯИ, 10-80-302, Дубна, 1988.
43. Закамский Л.А. и др. В кн.: Материалы семинара по обработке физической информации. Изд.ЕрФИ, Ереван, 1976, с.325.
44. Женеи М. В кн.: Материалы семинара по обработке физической информации. Изд.ЕрФИ, Ереван, 1976, с.237.
45. Байла И. и др. ОИЯИ, P10-80-162, Дубна, 1980.
46. Байла И., Ососков Г.А. ОИЯИ, P10-80-389, Дубна, 1980.
47. Байла И., Ососков Г.А. ОИЯИ, P10-80-237, Дубна, 1980.
48. Cittolin S., Taylor V.G., In: Proc. IEE Conf. on Video and Data Recording, Birmingham, 1976, p. 94.
49. Williams D.D. CERN DD/80/20, Geneva, 1980.
50. Байла И. и др. ОИЯИ, Б1, 10-82-II9, Дубна, 1982.
51. Байла и др. ОИЯИ, P10-82-653, Дубна, 1982.
52. Ососков Г.А. и др. ОИЯИ, P10-82-237, Дубна, 1982.
53. Бечер Ю. и др. В кн.: Материалы II Всесоюзного семинара по обработке физической информации, изд-во ЕрФИ, Ереван, 1977, с.428.
54. Дуда Р., Харт П. Распознавание образов и анализ сцен. "Мир", М., 1976.
55. Горелик А.Л., Скрипкин В.А. Методы распознавания, "Высшая школа", М., 1977.
56. Толкачев Д.А., Федотов О.П. В кн.: Материалы II семинара по обработке физической информации. Изд.ЕрФИ, Ереван, 1977, с. 277.
57. Prim R.C. Bell Sys. Tech.J., pp. 1389-1401, Nov. 1957.
58. Гудман С., Хидетниemi С. Введение в разработку и анализ алгоритмов, Мир, М., 1981.
59. Zahn C.T., IEEE Trans on Comp., v. C-20,1,1971, pp. 68-86.
60. Абдурахимов А.У. и др. ОИЯИ, P10-80-657, Дубна, 1980.
61. Алакоз А.В. и др. В кн.: Автоматизация научных исследований на основе применения ЭВМ (тезисы докладов VI Всесоюзной конференции). Изд. СО АН СССР, Новосибирск, 1981, с. 184.
62. Proc. of Conference on Computer Assisted Scanning, Padova, 1976.
63. Ананьева М.А. и др. ОИЯИ, 10-82-232, Дубна, 1982.
64. Anykeyev V.A. et al, In: Proc. of Conference on Computer Assisted Scanning, Padova, 1976, p. 207.
65. Астахов А.Я., Комов Г.М. ОИЯИ, P10-80-227, Дубна, 1980.
66. Bross A. et al, Nucl. Inst. Meth., v.166,3,1979, pp.367-378.

67. Кохонен Т. Ассоциативные запоминающиеся устройства, "Мир", М., 1982.
68. Королев Л.Н. Структуры ЭВМ и их математическое обеспечение, "Наука", М., 1978.
69. Тезисы Первого всесоюзного семинара по автоматизации научных исследований в ядерной физике и смежных областях. Изд-во "Дониш", Душанбе, 1980.
70. Proc. of the IEEE Computing Society on the Pattern Recognition, Chicago, IEEE Publishing Company, N.-Y., 1978.

Рукопись поступила в издательский отдел
24 марта 1983 года.