

**СООБЩЕНИЯ  
ОБЪЕДИНЕННОГО  
ИНСТИТУТА  
ЯДЕРНЫХ  
ИССЛЕДОВАНИЙ  
ДУБНА**

И 757

P10-86-666

Ц.И.Иоселиани\*, Г.А.Ососков, Н.И.Чернов,  
Е.Шандрикова

**СТАТИСТИЧЕСКИЙ ПОДХОД  
В ЗАДАЧЕ РАСПОЗНАВАНИЯ ОБРАЗОВ  
В ФИЗИКЕ ВЫСОКИХ ЭНЕРГИЙ**  
Точечная модель изображения треков

---

\*ИФВЭ Тбилисского государственного университета

**1986**

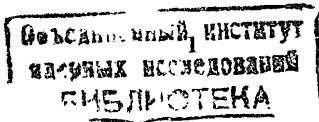
## Введение

Данная работа посвящена проблеме распознавания образов при анализе трековой информации в физике высоких энергий. Об актуальности проблемы свидетельствует большое число работ по этой тематике, представленных в обзорах /1-4/, трудах конференций /5-7/, а также в /8,9/. В них, в частности, отмечалось отсутствие достаточно общей формализованной теоретической модели трековых изображений /9/, существенное отличие распознавания таких изображений от классической теории распознавания /10,11/ и обосновывалась необходимость снижения огромной информационной емкости изображений (до  $10^8$  бит) путем введения различных систем целеуказаний и максимального использования априорных сведений. Целеуказания в фильмовой методике /12/ обеспечиваются путем предварительного отбора полезных снимков и измерения на каждом треке вручную одной или нескольких точек (маски). В случае одной точки маски говорят о минимальном наведении, а если она находится в вершине события, то - о наведении по вершине. В случае маски из нескольких точек - о полном наведении или о наведении по дорожке. Существуют достаточно надежные методы обработки данных при полном наведении /12/, и мы не будем далее их касаться.

Рост уровня целеуказаний ведет к упрощению задачи распознавания, но также к увеличению затрат на предварительный просмотр и измерение масок. Поэтому прилагались значительные усилия к разработке измерительных систем без наведения /12,13/. Однако эти системы ограничивались одним типом эксперимента, требовали высокой частоты появления событий на изображениях и больших вычислительных средств, т.к. основывались на технике прослеживания каждого из пучковых треков по всему кадру до конца трека или возможной вершины, после распознавания которой применялся метод минимального наведения.

В этой связи весьма многообещающими казались глобальные методы распознавания, основанные на применении теории графов (на что указывал П. Хафф еще в 1967 году /12/) или других способов кластеризации /4,8,14/, не получившие, однако, в то время широкого распространения при массовой обработке трековой информации из-за слишком больших требований к ресурсам времени и памяти ЭВМ.

В настоящей работе наметен новый, достаточно общий подход к преодолению трудностей в распознавании трековой информации, основанный на статистических моделях, рассчитанных на различные уровни целеуказа-



ний и способн съема информации. Подход основан на построении стохастической модели изображения и на программной реализации метода максимального правдоподобия для оценивания локальных параметров треков. Благодаря этому удалось по-новому решить проблему устойчивости оценок к влиянию шумового фона, загрязняющего выборку. В отличие от принятых в трековой обработке громоздких процедур выброса <sup>/4/</sup> или более тонких робастных методов оценивания <sup>/15,16/</sup> предлагаемый подход позволяет полностью учесть влияние шумовых точек, распределение которых включено в саму функцию правдоподобия. Возникающая при этом задача классификации точек решена с помощью сходящейся процедуры динамической кластерной перестройки <sup>/17/</sup>. Созданные алгоритмы, конечно, потребуют "доводки" на более обширном реальном материале, но уже полученные результаты показывают плодотворность предлагаемого статистического подхода при решении данного круга проблем.

### 1. Статистическая модель изображения события

След частицы в трековой камере, равномерно заполненной ионизируемой жидкостью или газом (пузырьковой или стримерной), образуется из пузырьков или стримеров, возникающих вокруг центров ионизации, распределенных по закону Пуассона <sup>/2,3,18/</sup>. Из последнего следует равномерность распределения какого-либо выбранного центра ионизации на фиксированном участке трека, а также показательное распределение длины промежутка между пузырьками или стримерами. Треки в искровых спектрометрах образуются искрами, появляющимися в заданных искровых промежутках, и случайность проявляется в пропадании искр при несрабатывании той или иной камеры <sup>/27/</sup>.

При дискретизации изображения в процессе сканирования с шагом  $h$  получаем набор отсчетов, т.е. точек  $\{p_i\}$ ,  $i=1, N$ , с координатами  $(x_i, y_i)$  в некотором прямоугольнике  $\Pi$  на плоскости (кадре). Информативная часть отсчетов образует треки-группы отсчетов, связанных близостью к гладким кривым, являющимся проекциями траекторий частиц на  $\Pi$ . Остальные отсчеты порождены шумовыми образованиями и сбоями сканатора. Близость отсчетов к соответствующим траекториям частиц, играющих роль осевых линий треков, характеризуется нормальным законом распределения отклонений  $d(p)$  отсчетов от линии трека:

$$f(p) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-d^2(p)/2\sigma^2). \quad (1)$$

В дискретизованных изображениях равномерность распределения отсчетов сохраняется, а распределение промежутков между отсчетами на треке становится геометрическим <sup>/23/</sup>, т.е. вероятность промежутка

из  $k$  скан-линий равна

$$P(\ell = k) = e^{-\lambda k} (1 - e^{-\lambda}). \quad (2)$$

Для пузырьковых и стримерных камер  $\lambda = gh$  ( $g$  - ионизационная плотность трека).

Заметим, что человек-оператор вне зависимости от знания физической стороны эксперимента обычно различает все треки и прослеживает их в шумах и взаимных пересечениях. Математическая формализация этой способности человека должна предполагать определенную идеализацию описания, необходимую для применения законов математики. В отношении предлагаемой модели это сводится к некоторым допущениям, обеспечивающим, главным образом, однородность условий измерения для каждого из треков.

Отметим, что изложенные ниже допущения вполне справедливы для автоматов с направленным сканированием, а в случае параллельного сканирования всего кадра - выполняются для основной части треков частиц больших энергий, направленных вдоль кадра.

1. Величины отклонений  $d(p)$  распределены согласно (1) с дисперсией  $\sigma^2$ , одинаковой для всех треков.

2. Вероятность  $w_j$  принадлежности отсчета  $p$  к  $j$ -му треку пропорциональна его длине  $L_j$ , вычисляемой как число скан-линий, пересекающих трек, т.е.  $w_j = \varepsilon L_j$ .

3. Вероятность появления отсчета  $p$ , относящегося к  $j$ -му треку на какой-либо из скан-линий, пересекающих этот трек, равна  $1/L_j$ .

4. Шумовые отсчеты распределены равномерно в  $\Pi$  с плотностью  $n(p) = n/S(\Pi)$ , где  $S(\Pi)$  - площадь  $\Pi$ .

5. Величины  $\sigma$ ,  $\varepsilon$  и  $n$  считаются известными или могут быть оценены независимо, т.е. являются характеристиками эксперимента, прибора, фильма, но не данной конкретной выборки.

В итоге распределение вероятностей каждой точки является смесью распределений, сосредоточенных вокруг каждого трека, и распределения шума. Плотность этого распределения может быть вычислена как

$$f(p) = \text{const} \left( \sum_{j=1}^M w_j f_j(p) \frac{1}{L_j} + n \right) = \frac{1}{Z} \left( \frac{\varepsilon}{\sigma\sqrt{2\pi}} \sum_{j=1}^M e^{-\frac{d_j^2(p)}{2\sigma^2}} + n \right), \quad (3)$$

где  $m$  - общее число треков,  $Z = (L_1 + \dots + L_M)\varepsilon$  - нормирующий множитель.

Поскольку точка  $p$  должна либо входить в какой-то трек, либо быть шумовой, то в сумме (3) одно из слагаемых будет, как правило, значительно превышать остальные. Поэтому можно заменить (3) на

$$f(p) = \begin{cases} \frac{\varepsilon}{z \sigma \sqrt{2\pi}} e^{-\frac{d_j^2(p)}{2\sigma^2}}, & \text{если } d_j^2(p) < D^2 = 2\sigma^2 \ln \frac{\varepsilon}{n \sigma \sqrt{2\pi}} \\ \frac{n}{z}, & \text{при одном единственном } j, j = \overline{1, M}, \\ & \text{если } d_j^2 \geq D^2 \text{ при всех } j. \end{cases} \quad (4)$$

**Замечание.** Если  $d_j^2 < D^2$  для нескольких  $j$ , то точка  $p$  находится в области пересечения или сближения треков. При этом возможна альтернатива: а) выбрать минимальное  $d_j(p)$  или б) исключить такие точки из рассмотрения, полагая  $f(p) = \frac{n}{z}$ . Вариант б) следует использовать в том случае, когда  $D$  меньше разрешающей способности прибора, вследствие чего данные из таких областей нельзя считать достоверными (рис. I).

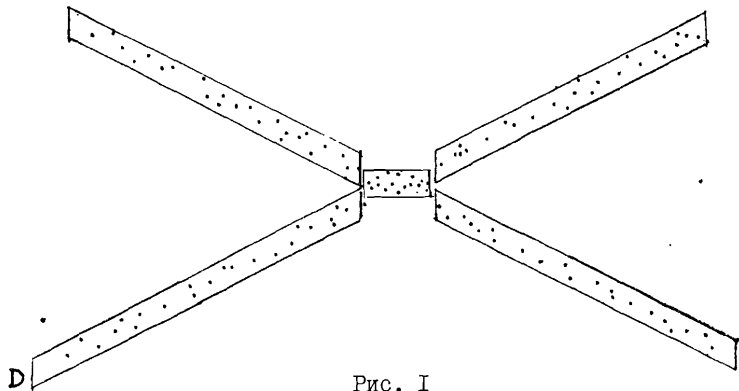


Рис. I

Формула (4) фактически задает классификацию точек на  $(M+1)$  группу, т.е. по кластерам, соответствующим  $m$  трекам, и кластеру шумовых точек.

## 2. Функция правдоподобия

Теперь мы можем приступить к выводу функции правдоподобия /19/ изображения. Совместную плотность распределения вероятностей для точек  $j$ -го трека получаем как

$$f(p_1, \dots, p_{N_j}) = \left( \frac{\varepsilon}{z \sigma \sqrt{2\pi}} \right)^{N_j} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{N_j} d_i^2(p_i)}, \quad j = \overline{1, M}, \quad (5)$$

а для шумовых точек

$$f(p_1, \dots, p_{N_0}) = \left( \frac{n}{z} \right)^{N_0}, \quad (6)$$

где  $N_1 + \dots + N_m + N_0 = n$  - общее число точек в  $\Pi$ . Перемножая плотности (5) для всех треков и (6), получаем полную логарифмическую функцию правдоподобия:

$$L(p_1, \dots, p_N) = -(N - N_0) \ln \frac{\varepsilon}{z \sigma \sqrt{2\pi}} - N_0 \ln \frac{n}{z} + \quad (7)$$

+  $\frac{1}{2\sigma^2} \sum_{j=1}^M \sum_{i=1}^{N_j} d_i^2(p_i^{(j)}) + N \ln z$ ,  
где  $p_1^{(j)}, \dots, p_{N_j}^{(j)}$  - точки  $j$ -го кластера (трека).

Минимизация (7) эквивалентна решению задачи

$$\tilde{L}(p_1, \dots, p_N) = N_0 D^2 + \sum_{j=1}^M \sum_{i=1}^{N_j} d_i^2(p_i^{(j)}) \rightarrow \inf \quad (8)$$

с  $D^2$ , известным из предположения (5) (здесь мы пренебрегаем зависимостью множителя  $z$  от  $L_1, \dots, L_M$  - об оценивании величин  $L_i$  речь пойдет ниже).

Если записать уравнение осевой линии  $j$ -го трека в виде  $y = g_j(x, \bar{\theta}_j)$  ( $\bar{\theta}_j$  - вектор параметров), то отклонение точки  $(x_i, y_i)$  от этого трека имеет вид

$$d_j(p_i^{(j)}) = y_i - g_j(x_i, \bar{\theta}_j), \quad (9)$$

так что искомая оценка максимального правдоподобия для параметров всех  $m$  треков может быть получена минимизацией (8) по всем возможным разбиениям множества измеренных точек на  $(M+1)$  подмножество, а для каждого из подмножеств, соответствующих трекам, - по параметрам этих треков.

Следует сразу же отметить ряд существенных затруднений в применении (8) для решения проблемы распознавания треков, которые, возможно, явились причиной столь малого внимания многочисленных исследователей этой актуальной проблемы к использованию метода максимального правдоподобия, обычно наиболее привлекательного с точки зрения качества получаемых оценок.

Прежде всего, вид формулы (8) не дает явного конструктивного подхода для группировки точек по трекам, более того, при неизвестном  $M$  минимизация (8) и (7) приводит нас в тупик, ловушку, т.к. минимум достигается в тривиальном решении  $m = n$ , когда число треков становится равным числу всех точек.

Кроме того, определенные трудности вызывает вычисление величин  $d_j(p)$ , т.к. уравнение трека даже в идеальном случае однородного

магнитного поля и отсутствия дисторсий регистрирующего оборудования является сложной кривой четвертого порядка.

Тем не менее анализ этих затруднений показывает, что они преодолимы. Начиная с последнего из них, укажем на целесообразность применения аппроксимационных моделей треков, позволяющих вести распознавание по "сырым" данным без использования калибровочных преобразований и учета неоднородностей магнитного поля. Для аппроксимаций использовались различные виды кривых: параболы /20/, кубические сплайны /21/, окружности /4, 22/. Применение последней, как показала практика, наиболее надежно, а при наличии быстрых алгоритмов подгонки /23/ не уступает остальным и в скорости вычислений. В такой локально-циркулярной модели трека величины  $a_j(p)$  могут, например, вычисляться как расстояния от точки  $p$  до окружности, получаемой подгонкой по точкам  $p'$  трека, попавшим в  $r$ -окрестность точки  $p$  ( $|p' - p| < r$ ). Таким образом, в дальнейшем мы будем предполагать, что распознаваемые треки достаточно близки к окружностям. Отметим в этой связи полезность введения новой параметризации окружности, снимающей вычислительные проблемы при оценке радиуса для треков малой кривизны (см. приложение).

### 3. Динамическая кластеризация

Если число треков известно априори (как при распознавании с минимальным наведением), то задача (8) решается методом, аналогичным методу динамической кластерной перестройки /17/, алгоритм которого (АПК1) выглядит следующим образом:

**Шаг 0.** Выбирается начальная кластеризация по  $(m+1)$  группе (исходя из априорной информации и любых методов инициализации поиска треков /4/).

**Шаг 1.** Для каждой точки  $p$  вычисляется расстояние  $a_j(p)$  до каждого трека и производится перестройка кластеров по правилу (4).

**Шаг 2.** Если состав кластеров после шага 1 изменился, то возвращаемся к шагу 1, в противном случае - заканчиваем работу.

При использовании варианта (а) в замечании к формуле (4) сходимость алгоритма обеспечена в силу того, что значение  $L$  уменьшается после каждого выполнения шага 1. При использовании варианта (б) алгоритм не обязательно сходится (и существуют весьма искусственные примеры расходимости), но проверка на имитационных моделях трековой информации дает хорошие результаты и в этом случае (см. ниже).

При отсутствии сведений о числе треков (как, например, при распознавании без наведения и без предварительного просмотра) обобщение АПК1 позволяет в принципе решить и эту задачу. Тип изучаемого физи-

ческого взаимодействия определяет минимально возможное количество вторичных частиц, т.е.  $m_1$  - минимально возможное число треков. С помощью АПК1 можно выделить на оцифрованном изображении  $M_1$  треков. Если  $m > m_1$ , то недостающие треки останутся в шумовом кластере, к которому можно снова применить АПК1 для "отщепления" еще одного трека и, подобрав подходящий эмпирический критерий минимально допустимого трека, пытаться повторять эту процедуру до исчерпания.

Очевидно, это не самый быстрый и надежный путь, поэтому вернемся к задаче (8) при неизвестном  $m$ . Ловушка, о которой говорилось, вызвана неполнотой функции плотности (5), где не учтен тот факт, что некоторые точки траектории не обнаружены на треке. Для его полного описания следует включить в  $f(p_1, \dots, p_N)$  распределение пробелов  $l_1, \dots, l_{N-1}$  между точками, а также расстояние от видимых концов до выхода трека за края кадра ( $l_0$  и  $l_{N_j}$ ). С учетом независимости  $l_i$  и их распределения (2) получим вместо (5)

$$f^*(p_1^{(j)}, \dots, p_N^{(j)}) = \left( \frac{\epsilon}{z \sigma \sqrt{2\pi}} \right)^{N_j} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{N_j} a_i^2(p_i^{(j)})} \cdot e^{-\lambda \sum_{i=0}^{N_j} l_i} (1 - e^{-\lambda})$$

Величин  $l_0$  и  $l_{N_j}$  мы не знаем, поэтому предположим, что  $l_0 + l_{N_j} = l$  - некоторая константа, подбираемая экспериментально, как  $\sigma$  и  $\epsilon$ .

Формула (7) может быть переписана в виде

$$L^*(p_1, \dots, p_N) = L(p_1, \dots, p_N) + \lambda m l + \lambda \sum_{j=1}^m l^{(j)} + \ln(1 - e^{-\lambda})$$

где  $l^{(j)} = l_1^{(j)} + \dots + l_{N_j-1}^{(j)}$ . Формула (8) примет вид

$$\tilde{L}^*(p_1, \dots, p_N) = N_0 D^2 + \sum_{j=1}^M \sum_{i=1}^{N_j} a_j^2(p_i^{(j)}) + \zeta m + \lambda' \sum_{j=1}^M l^{(j)} \quad (10)$$

где  $\lambda' = 2\sigma^2 \lambda$ ,  $\zeta = 2\sigma^2 \lambda l$ .

Появление члена  $\zeta m$ , выполняющего роль "штрафа" за попытку чрезмерного увеличения числа треков, компенсирует уменьшение остальных слагаемых при минимизации  $\tilde{L}^*$ .

Алгоритм перестройки кластеров при неизвестном  $m$  (АПК2) получается из АПК1 введением дополнительных шагов.

**Шаг 3.** Каждый кластер пытаемся разбить на два кластера. Если при этом  $\tilde{L}^*$  уменьшается, то фиксируем разбиение (увеличиваем  $m$  на единицу) и переходим к шагу 2.

**Шаг 4.** Каждую пару кластеров пытаемся объединить в один кластер. Если при этом  $\tilde{L}^*$  уменьшается, то фиксируем слияние, уменьшаем на 1 число кластеров и возвращаемся к шагу 1, если нет - заканчиваем работу.

В шаге 2, соответственно, вместо окончания работы надо вставить переход к шагу 3.

Проблема сходимости АПК2, так же, как и АПК1, решается благодаря уменьшению  $\hat{\gamma}^*$  на каждом шаге. Начальное значение  $m$  выбирается на основе априорной информации или путем назначения некоторой минимальной величины (например,  $m=0$ ).

Для проверки работы АПК2 смоделировано несколько сотен изображений взаимодействия с одним пучковым треком и несколькими вторичными (от 5 до 10) с кривизной разного знака и величины. Угол вылета разгравлялся равномерно в конусе  $\pm 30^\circ$  от направления пучка. Уровень шумов составлял 50% от общего числа точек на треках. Примеры смоделированных изображений приведены на рис.2.

Из методических соображений алгоритм АПК2 применялся без каких-либо предварительных целеуказаний. Начальная кластеризация выбиралась с  $m=0$  (все точки - в шумовой кластер), "отщепление" треков от шумового кластера производилось путем случайного выбора трех точек в нем и проведения через них дуги окружности с занесением в новый трек всех точек из ее  $D$  - окрестности (величина  $D$  вычислялась согласно формуле (4)).

Эффективность алгоритма АПК2 можно оценить с помощью вероятностей  $\alpha$  - необнаружения одного из истинных треков и  $\beta$  - ложного обнаружения. В результате моделирования были вычислены доверительные границы для  $\alpha$  и  $\beta$  при 95% уровне значимости:  $\alpha < 1,9 \cdot 10^{-2}$ ,  $\beta < 1,2 \cdot 10^{-2}$ .

Заметим, что алгоритм АПК2 с локальной аппроксимацией трека дугами окружности способен распознать треки переменной кривизны и даже спиралевидные треки. Иллюстрирующее изображение на рис.3 было распознано алгоритмом за 2 мин на ЭВМ ЕС-1061.

**Выводы.** 1. Глобальные алгоритмы АПК1 и АПК2 показали очень высокую эффективность на модельных данных значительной сложности и зашумленности.

2. Два взаимосвязанных фактора препятствуют прямому приложению этих алгоритмов при массовой обработке реальных трековых данных:

- чрезмерные временные затраты, явившиеся следствием отказа от целеуказаний и применения метода "грубой силы", т.е. простого случайного поиска при формировании начальных кластеров и их перестройке;
- отсутствие средств учета специфики обрабатываемого изображения, присущей конкретному типу эксперимента, способу сканирования, статистике пробелов и т.д.

3. Необходимо исследование по применению как целеуказаний, учитывающих априорные сведения об изображении, так и методов сокращения перебора (проверка кандидатов в треки на гладкость, отсутствие изломов, больших разрывов и т.д.), а также по выбору оптимальной процедуры

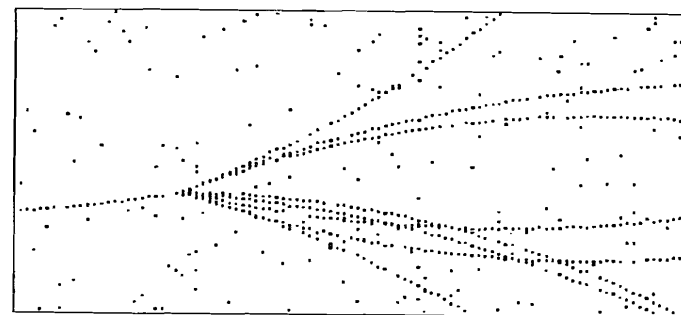
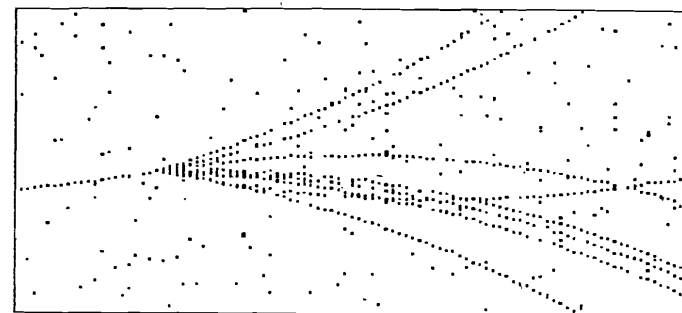


Рис. 2

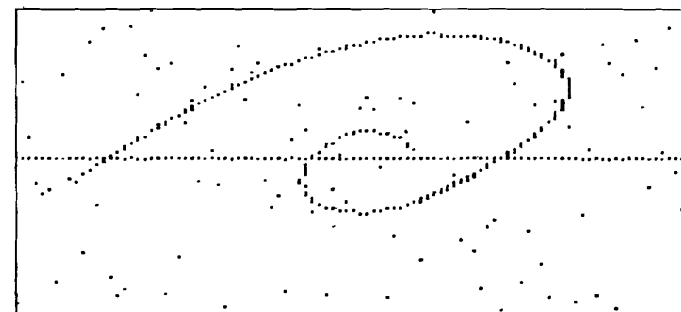


Рис. 3

перестройки (например, можно на шаге 3 проводить не одно, а все разбиения, уменьшающие функционал (10), или же только те, которые уменьшают его максимально).

Кроме того, реализация алгоритмов должна учитывать тип используемой ЭВМ, состав ее математического обеспечения.

От всего комплекса мер, направленных на ускорение и оптимизацию алгоритмов, можно ожидать многократного повышения их эффективности.

4. В своем настоящем виде АПКЗ не ориентирован на использование его с данными предварительного сжатия трековой информации в виде линейных трек-элементов. Модификации АПКЗ, учитывающие такой характер данных, являются нетривиальными. Планируется рассмотреть их в следующей работе.

#### Приложение

Общепринятая параметризация окружности  $(x-a)^2 + (y-b)^2 = R^2$  становится неудобной при аппроксимации треков малой кривизны, т.к. параметры  $a$ ,  $b$ ,  $R$  могут неограниченно возрастать и это затрудняет вычисление на ЭВМ. В действительности нам необходимы не значения  $a$ ,  $b$ ,  $R$ , а формула для вычисления расстояния от произвольной точки трека  $p=(x,y)$  до аппроксимирующей его дуги окружности:

$$d(p) = \left| \sqrt{(x-a)^2 + (y-b)^2} - R \right|.$$

Для треков  $d(p) \ll R$ , поэтому (см. также /23/)

$$d(p) = \left| \frac{(x-a)^2 + (y-b)^2 - R^2}{\sqrt{(x-a)^2 + (y-b)^2 + R}} \right| \approx \left| \frac{(x-a)^2 + (y-b)^2 - R^2}{2R} \right|.$$

Полученное выражение позволяет перейти к новым параметрам:

$$d(p) \approx \frac{1}{2R}(x^2 + y^2) - \frac{a}{R}x^2 - \frac{b}{R}y^2 + \frac{a^2 + b^2 - R^2}{2R} = \alpha(x^2 + y^2) + \alpha x + \beta y + \gamma,$$

где  $\alpha = (2R)^{-1}$ ,  $\alpha = -a/R$ ,  $\beta = -b/R$ ,  $\gamma = (a^2 + b^2 - R^2)(2R)^{-1}$  — новые параметры окружности, не имеющие особенностей для дуг малой кривизны. Для интерпретации этих параметров рассмотрим систему координат, связанную с центром тяжести системы точек трека. В этой системе координат  $\alpha x + \beta y + \gamma = 0$  есть уравнение касательной к аппроксимирующей дуге окружности в середине трека (рис.4), а  $2\alpha$  — кривизна дуги. Заметим, что параметры  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\alpha$  связаны между собой соотношением  $\alpha^2 + \beta^2 = 4\alpha\gamma + 1$ . Их оценка проводится в точности тем же методом, что и оценка  $a$ ,  $b$ ,  $R$  в /23/, за исключением самых

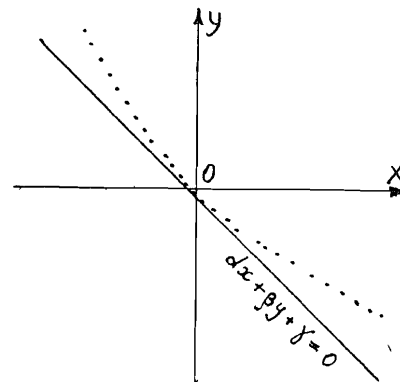


Рис. 4

последних формул, где вычисления  $a$ ,  $b$ ,  $R$  заменяются на непосредственные вычисления  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\alpha$ , позволяющие избежать малых знаменателей в случае треков малой кривизны.

#### Литература

1. Автоматический анализ сложных изображений, сб. под ред. Э.М.Бравермана. Мир, М., 1969.
2. Автоматическая обработка данных с пузырьковых и искровых камер, сб. под ред. Б.С.Розова. Атомиздат, М., 1971.
3. Стрэнд Р. Распознавание оптических образов при экспериментах на трековых камерах с частицами высоких энергий, в сб.: Распознавание образов при помощи цифровых вычислительных машин, под ред. Л.Хармона. Мир, М., 1974, с. 15-37.
4. Ососков Г.А. Применение методов распознавания образов в физике высоких энергий. ОИЯИ, П10-83-187, Дубна, 1983.
5. Труды Международного симпозиума по вопросам автоматизации обработки данных с пузырьковых и искровых камер. ОИЯИ, Д10-II-6142, Дубна, 1972.
6. Материалы II Всесоюзного совещания по обработке физической информации. Изд. ЕрФИ, Ереван, 1978.
7. Труды V Международного совещания по проблемам математического моделирования, программирования и математическим методам решения физических задач. ОИЯИ, Д10, II-84-818, Дубна, 1984.
8. Zahn С.Т. Graph-theoretical methods for detecting and describing gestalt clusters. IEEE Trans. on Comp., v.C-20, No1, 1971, p.68-86.
9. Бережной В.А., Жигунов В.П., Никитин С.Т. Математические аспекты распознавания снимков с пузырьковых камер. ИФВЭ, 80-86, Серпухов, 1980.
10. Ту Дж., Гонзалес Р. Принципы распознавания образов. Мир, М., 1978.

11. Дуда Р., Харт П. Распознавание образов и анализ схем. Мир, М., 1974.
12. White H.S., Hall D. Performance of the DARP System. Proc. Int. Conf. on Data Handling Systems in High-Energy Physics. 1970, CERN 70-21, Geneva, p.761-803.
13. Говорун Н.Н., Дикусар Н.Д. Автоматическая обработка изображений на снимках с магнитного искрового спектрометра. ОИЯИ, IO-IO33I, Дубна, 1976.
14. Ососков Г.А., Пахомов В.М. В кн.: Совещание по программированию и математическим методам решения физических задач. ОИЯИ, ДПО, II-II264, Дубна, 1978, с.288.
15. Хьюбер П. Робастность в статистике. Мир, М., 1984.
16. Куняев С.В., Ососков Г.А., Чернов Н.И. Статистические методы распознавания и робастного оценивания параметров треков. ОИЯИ, P10-84-553, Дубна, 1984.
17. Diday E., Simon J. Clustering analysis. In: Digital Pattern Recognition, ser.: Communication and Cybernetics, No10, Springer-Verlag, Berlin, 1984, p. 47-94.
18. Ососков Г.А. Теория вероятн. и ее применение. Т. I, вып. 2, 1956, с. 274-282.
19. Крамер Г. Математические методы статистики. ИЛ, М., 1948.
20. Деарт А.Ф., Кистенев Э.П. и др. Труды РТИ АН СССР, № IO, М., 1972, с. 6-12.
21. Тилипман М.Ш. В кн.: Материалы семинара по обработке физической информации. Изд. ЕрПИ, Ереван, 1976, с. 314-331.
22. Dunn L.A. In: Proc. of the Conf. on Computer Assisted Scanning, Padova, 1976.
23. Ососков Г.А., Чернов Н.И. ОИЯИ, P10-84-7, Дубна, 1984.

Рукопись поступила в издательский отдел  
8 октября 1986 года.

Иоселиани Ц.И. и др.

P10-86-666

Статистический подход в задаче распознавания образов в физике высоких энергий. Точечная модель изображения треков

Рассматривается задача распознавания треков элементарных частиц среди двумерного множества точек, полученных при оцифровке изображения. Предложена формализованная математическая модель этой задачи, в которой каждый трек предполагается близким к окружности, а число треков также подлежит определению. Показано, что оценка максимального правдоподобия для параметров всех треков одновременно в сочетании с методом динамической кластерной перестройки дает решение задачи даже в отсутствие предварительных целеуказаний. Полученный алгоритм распознавания был проверен на монте-карловской модели данных с водородной пузырьковой камеры и показал высокую эффективность.

Работа выполнена в Лаборатории вычислительной техники и автоматизации ОИЯИ.

Сообщение Объединенного института ядерных исследований. Дубна 1986

Перевод О.С.Виноградовой

Ioseliani C.I. et al.

P10-86-666

Statistical Approach in Some Problem of Pattern Recognition in High Energy Physics. Pointed Model of Track Image

The problem of particle track recognition for a two-dimensional data set of points obtained by a digitalization of image is considered. For this problem a formal mathematical model of data is suggested. The shape of each track is supposed to be an arc of a circle or close to it. The number of tracks is to be found too. It is shown that maximum - likelihood estimation of parameters of all tracks simultaneously combined with dynamical cluster method solves this problem under zero-guidance conditions. The algorithm of recognition has been tested by the Monte Carlo data model of a hydrogen bubble chamber. Its high efficiency is shown.

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

Communication of the Joint Institute for Nuclear Research. Dubna 1986