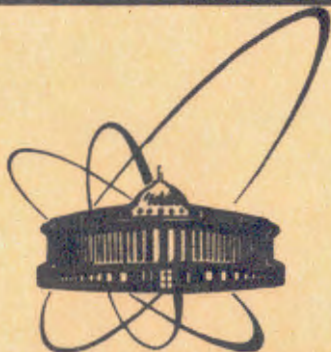


85-516



СООБЩЕНИЯ
ОБЪЕДИНЕННОГО
ИНСТИТУТА
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ
ДУБНА

УДЧО В
5726/85

P10-85-516

Я.Балгансурен, Н.А.Буздавина, В.В.Глаголев,
Н.Н.Говорун, А.Дирнер, Л.Дорж, Т.И.Забой,
А.Г.Заикина, В.Г.Иванов, А.П.Кретов, В.П.Миролюбов,
В.В.Первушов, И.И.Шелонцев, А.А.Хачатрян

ВЫСОКОАВТОМАТИЗИРОВАННАЯ СИСТЕМА
МАТЕМАТИЧЕСКОЙ ОБРАБОТКИ
ФИЛЬМОВОЙ ИНФОРМАЦИИ

1985

В ходе проводимых в ОИЯИ работ по повышению уровня автоматизации процесса обработки фильмовой информации была разработана устойчивая к внешним сбоям система математической обработки экспериментальных данных, обеспечивающая автоматический подсчет данных по заданным цепочкам программ, разбор и предварительный анализ результатов счета.

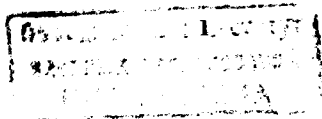
Испытания системы проводились в ходе обработки результатов *αp*-эксперимента^{/1/} и показали, что ее применение обеспечивает стабильную и практически безаварийную обработку данных, резкое сокращение затрат ручного труда и сокращение сроков обработки в 3-4 раза. Простота "общения" пользователя с системой позволяет работать с ней без специальной подготовки.

Данная работа посвящена общему рассмотрению системы, перечню решенных в ходе ее создания задач и результатам ее испытаний в *αp*-эксперименте.

§ I. Основные трудности, возникающие в процессе обработки фильмовой информации традиционными методами

Принципиальная схема процесса обработки фильмовой информации в рамках традиционной методики показана на рис. I. В качестве измерительного прибора здесь используется сканирующий автомат типа НРД .

Традиционные методы обработки фильмовой информации сложились около двадцати лет тому назад, когда быстродействие ЭВМ было относительно мало, а в качестве носителей информации использовались магнитные ленты и частично перфокарты^{/2/}. Поэтому ЭВМ в системах обработки использовались в основном для управления работой просмотрово-измерительных устройств и проведения трудоемких вычислений. Что касается разбора и анализа результатов счета, а также управления всем этим достаточно сложным процессом, то все эти операции в основном производились людьми. Для отбора данных на ленты суммарных результатов (ЛСР) эксперимента использовались перфокарты, на которых задавалась необходимая для их идентификации информация. Вследствие этого процесс обработки был многоступенчатым и требовал больших затрат ручного труда.



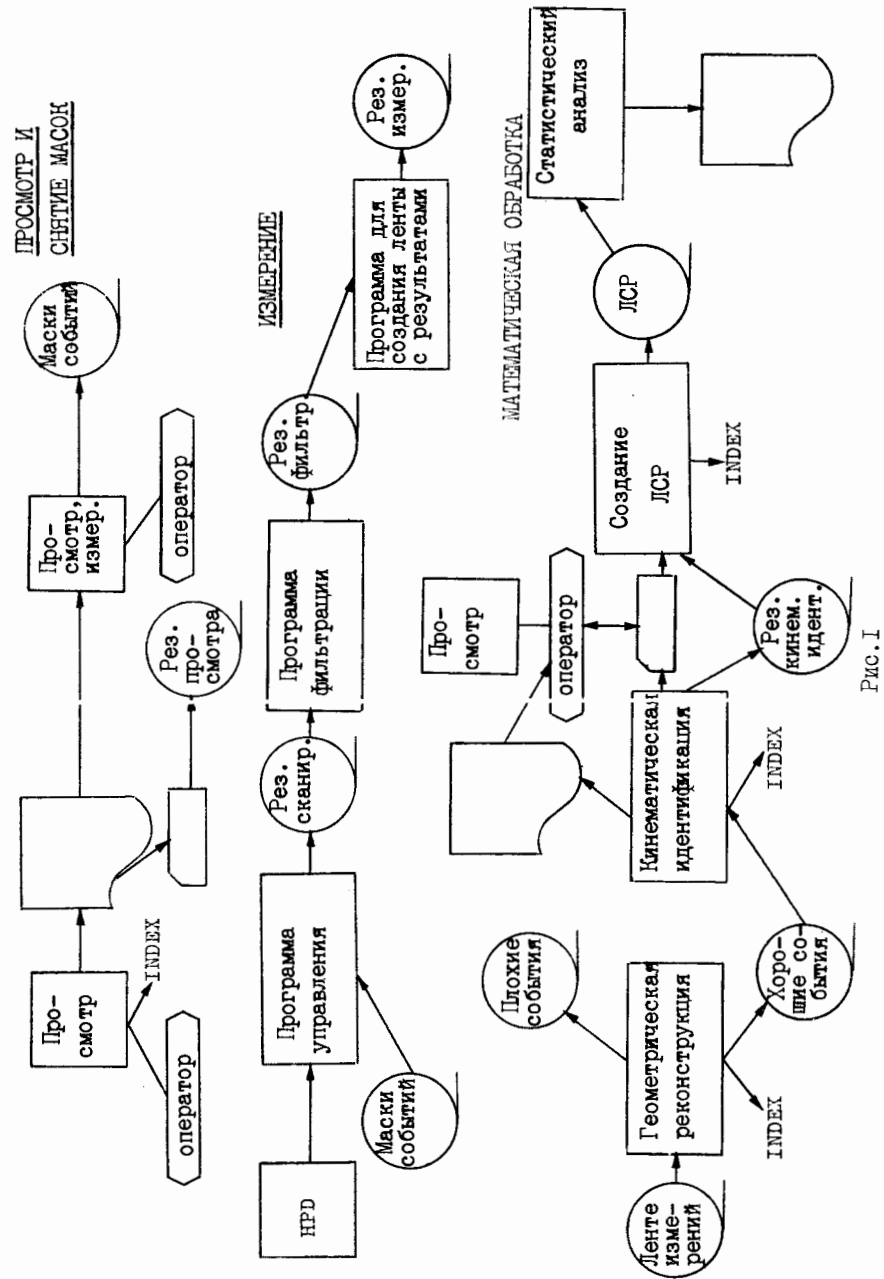


Рис. 1

Повышение быстродействия ЭВМ и производительности измерительных устройств привело к тому, что реальные скорости обработки стали существенно отставать от потенциальных возможностей имеющейся аппаратуры и вычислительных машин. Так, например, в условиях ОИЯИ для измерения 30 тыс. событий сравнительно простой топологии с жидководородной камеры требуется не более 15 недель работы НРД и около 100 часов времени ЭВМ CDC-6500. Что касается последующего анализа результатов, то он занимает несколько лет, т.к. один физик за неделю работы может просмотреть не более 250-300 событий^{/3/}. По результатам просмотра составляются различного рода списки, отображаются перфокарты, часть событий отправляется на повторные измерения и т.д. и т.п. Все это, естественно, требует времени, ровно как и ожидание повторных измерений, которые нужно снова анализировать.

Таким образом, сама традиционная организация процесса обработки filmовой информации, в основном ориентированная на активное участие в ней человека и не использующая многие возможности современных ЭВМ, является одним из серьезнейших препятствий на пути его автоматизации^{/4/}.

Другой ряд трудностей связан с тем, что программы анализа filmовой информации имеют весьма низкий приоритет на современных ЭВМ, которые работают в мультипрограммном режиме. Эти трудности обусловлены как размером основных программ обработки, которые занимают много места в оперативной памяти ЭВМ и тем самым снижают эффективность работы центральных процессоров, так и организацией программ, которую, вообще говоря, можно и улучшить. Рассмотрим в связи с этим вопросы повышения эффективности работы программ с точки зрения мультипрограммной ЭВМ.

Как следует из рисунка 1, в качестве основного носителя информации в системе используется магнитная лента. Имеющиеся на ЭВМ CDC-6500 магнитофоны позволяют записывать на одну магнитную ленту информацию общим объемом 19,2 Мб (2,5 млн. слов). На лентах исходных данных находится от одной до четырех тысяч измеренных событий, а ленты с результатами геометрической реконструкции и кинематической идентификации содержат около тысячи событий. Поскольку обсчет результатов измерений камерных фотографий производится в наших условиях порциями по несколько сот событий в каждой, то одни и те же магнитные ленты используются в нескольких сеансах счета, что приводит к аварийным остановкам работы программ из-за не очень надежной работы магнитофонов. Кроме того, при каждом повторном сеансе счета необходимо задавать программам количество событий, которые нужно пропускать на лентах исходных данных и результатов. Опыт работы с магнитными лентами на ЭВМ CDC-6500 показал, что для обеспечения надежности рабо-

ты программ и повышения эффективности использования центрального процессора целесообразнее производить считывание исходных данных и запись результатов счета, используя магнитный диск^{15/}.

Информация, необходимая для физической интерпретации одного события, после ряда проведенных сокращений занимает около тридцати строк машинной выдачи. И это обстоятельство накладывает определенные ограничения на число событий, которые можно обсчитать в одном сеансе. Дело в том, что количество информации, выдаваемой задачей на печать, ограничено. Поэтому в одном сеансе может обрабатываться не более 250-300 событий, а обычно эта цифра порядка ста. Таким образом, сокращение выдаваемой на печать информации, несмотря на уже принятые меры, все еще остается актуальной задачей.

Много сложностей создает использование для переноса данных о физических гипотезах специальных перфокарт, которые нужно получить, проинтерпретировать, разобрать и снова ввести в ЭВМ. Поскольку здесь речь идет об обработке и вводе в ЭВМ десятков тысяч перфокарт, то вероятность потерь на этом этапе достаточно велика, и это обстоятельство также чрезвычайно усложняет процедуру формирования лент суммарных результатов.

Крайне важно в процессе обработки иметь возможность оперативно получать информацию о состоянии дел на данный момент. Для этого в системах предусмотрено создание специальных файлов каталогов, хранящих "историю обработки" каждого события. Содержимое этих файлов обновляется после каждого сеанса счета с помощью специальных программ.

Современные системы математической обработки фильмовой информации это большие и сложные человеко-машинные комплексы, в работе которых участвуют десятки людей различных специальностей и квалификаций. Надежная работа такого типа систем невозможна без оперативного контроля как исходных данных, так и результатов счета, получаемых на каждом из основных этапов процесса обработки. Для решения этой задачи в наших условиях были созданы специальные программы и разработана концепция программного сопровождения. Суть последней заключается в том, что получаемые на каждом этапе процесса обработки результаты, начиная с ввода данных, анализируются специальными программами контроля, которые проверяют достоверность получаемых результатов и прекращают дальнейшую обработку, если полученные данные не удовлетворяют заданным критериям.

Таким образом, ведение каталога обрабатываемых событий и организация контроля результатов требуют увеличения числа программ, по которым ведется обсчет данных. Это обстоятельство также создает определенные затруднения в работе, так как чем больше требуется выходов на машину, тем медленнее идет процесс обработки. Эти затруднения

можно устранить, организовав автоматический обсчет данных по заданной цепочке программ с помощью специальной управляющей программы или ограниченного числа процедур, написанных на специальном языке ССЛ^{16/}

Анализ процесса обработки фильмовой информации показывает, что основным препятствием на пути его автоматизации является существующая практика физической интерпретации событий и формирования ЛСР. Именно эти операции являются наиболее "узкими местами" всего процесса и требуют максимальных затрат ручного труда. Дело осложняется еще и тем, что заметная часть (5-30%) измеренных событий бракуется на различных этапах процесса обработки. Для того, чтобы избежать появления в экспериментальных результатах каких-либо систематических ошибок приходится анализировать причины отказов в каждом конкретном случае, а часть событий перемерять по несколько раз. Все это чрезвычайно усложняет процедуры учета и требует дополнительных затрат ручного труда. Сокращение последних может быть достигнуто за счет более активного использования ЭВМ на указанных операциях и разработки программ, заменяющих человека на тех операциях, которые можно алгоритмизировать.

Следует также отметить слабую защищенность систем математической обработки фильмовой информации от внешних помех, под которыми здесь понимаются фатальные для работы программ ошибки в исходных данных (результатах измерений камерных фотографий), сбои в работе ЭВМ, ошибки персонала и, наконец, исследовательский характер работ, требующий внесения изменений в уже отмеченный процесс по ходу обработки.

§ 2. Общая организация предлагаемой системы обработки фильмовой информации

Принципиальная схема процесса обработки фильмовой информации в рамках предлагаемой системы показана на рис.2.

На этапе просмотра производится также визуальная оценка плотности почернения вторичных треков событий путем их сравнения с пучковым треком. Результаты просмотра заносятся на перфокарты, а затем записываются в виде перманентного файла (ФРП) на магнитный диск ЭВМ CDC-6500. Содержащаяся в этом файле информация используется затем для составления списков измеряемых событий с указанием типа прибора, на котором их нужно производить. Таким образом, результат работы этого этапа - файл результатов просмотра, содержащий наряду с обычной информацией также данные о плотностях почернения каждого вторичного трека событий.

Составленные с помощью ФРП списки событий для измерений на полуавтоматах и сканирующем автомате типа НРД передаются в соответствующую

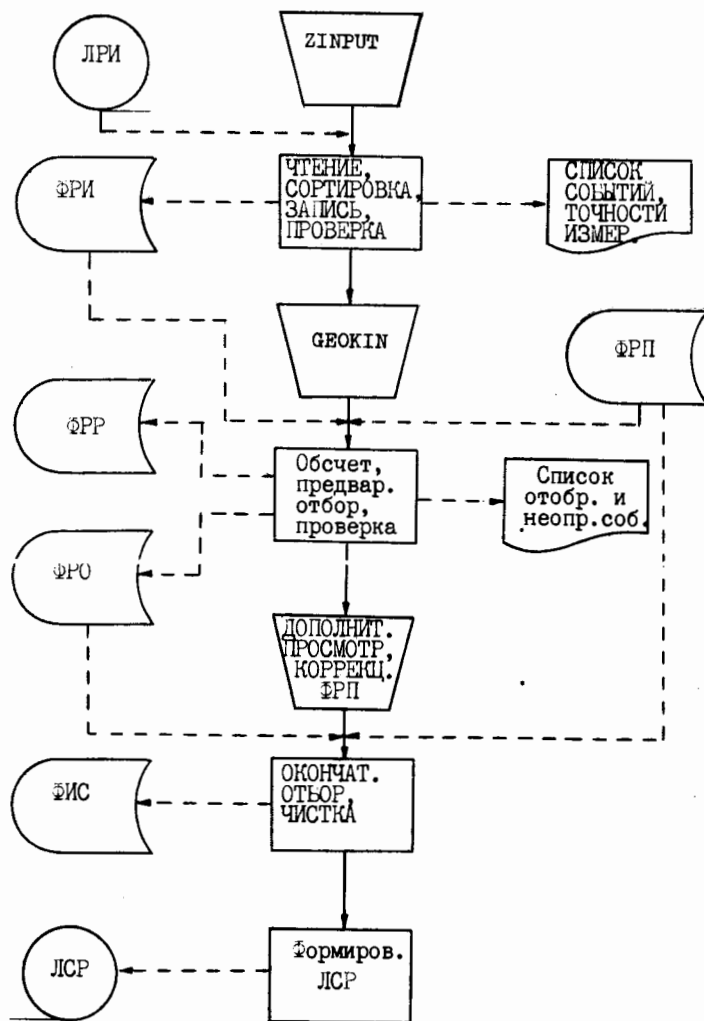


Рис.2

щий отдел, из которого физики получают магнитные ленты с результатами обмена элементов событий (ЛРИ). Поскольку измерения фотопленок производятся на трех различных измерительных устройствах и предварительная обработка данных и комплектация событий производятся на различных ЭВМ, то после измерений имеются ЛРИ трех типов.

На следующем этапе содержимое ЛРИ любого типа переписывается на магнитный диск, события упорядочиваются по номерам фотопленок и стереоскадров, информация преобразуется в стандартную форму, и производится проверка точностей измерений различных элементов событий на их стереоснимках. Все операции этого этапа выполняются в рамках одного задания ЭВМ, оформленного в виде процедуры (ZINPUT) языка SCL.

Затем производится восстановление пространственной картины измеренных событий, кинематическая идентификация результатов реконструкции, предварительный отбор физических гипотез по заданным критериям и, наконец, проверка получаемых результатов. Все эти операции также выполняются в рамках одного задания ЭВМ с помощью процедуры GEOKIN. Поскольку массив исходных данных, содержащихся на одной магнитной ленте, обрабатывается за несколько сеансов счета, то программы этой процедуры (GEOKIN и FIG) сами находят файлы исходных данных (ФРИ, ФРР), результатов счета (ФРР, ФРО) и число записей на каждом из них, которые нужно пропустить в каждом из последующих сеансов. Иначе говоря, программы системы автоматически находят число записей на файле исходных данных и сами объединяют результаты счета каждого двух последующих сеансов счета в один файл результатов без вмешательства человека до тех пор, пока размеры файлов не достигнут предельных значений. Тогда с помощью специальной процедуры DUMP производится сброс файлов результатов на магнитные ленты для хранения и последующего использования.

Процедура DUMP рассчитана на два режима сброса: промежуточный и окончательный. Первый предназначен для сброса файлов результатов программ GEOKIN и FIG на промежуточном этапе, когда размеры файлов достигли максимальной величины, а файл исходных данных еще не исчерпан. В этом случае после сброса файлов результатов на магнитные ленты, из файла результатов программы GEOKIN выбираются головные рекорды и из них формируется новый файл результатов. Эта операция нужна для того, чтобы программа GEOKIN могла определить число событий, которые ей нужно пропустить на файле исходных данных при следующих сеансах счета. И только после этого производится уничтожение старого файла результатов программы GEOKIN. При окончательном сбросе файлов результатов, когда весь массив исходных данных, содержащихся на ЛРИ, обработан, на магнитные ленты копируются оба файла результатов, после чего они уничтожаются, и система готова к анализу следующей ЛРИ.

Для того, чтобы в максимальной степени сократить затраты ручного труда на разбор и анализ результатов счета перед копированием файлов на ленты по результатам счета программы FIG составляется паспорт обработанного массива данных. В него входят следующие списки:

- список хорошо измеренных событий;
- список дважды измеренных и оба раза забракованных событий, которые следует перевести в разряд неизмеримых;
- список событий, измеренных дважды;
- список забракованных событий, которые нужно перемерить.

После этого производится отбор физических гипотез с учетом визуальных оценок ионизаций треков, проведенных при предварительном просмотре. Эта операция выполняется в два этапа.

Сначала на основе сопоставления результатов реконструкции и просмотра событий выделяются события, для которых результаты программной идентификации не вызывает сомнений. Для остальных на печать выдается информация как о результатах просмотра, так и идентификации с тем, чтобы экспериментаторы могли просмотреть неоднозначно идентифицированные события и уточнить критерии их отбора. Что касается однозначно идентифицированных событий, то для них на печать выдаются лишь номера отобранных гипотез и результаты визуальных оценок ионизаций треков. Эти данные позволяют экспериментаторам производить выборочный, а в некоторых случаях и полный контроль результатов просмотра и сохраняют за ними возможность участия в отборе гипотез для каждого события. На этом этапе также имеется возможность проконтролировать работу просмотрщиков и в случае необходимости внести изменения в ФРП.

После проверки и исправления результатов просмотра программа отбора запускается в режиме переписи физических гипотез на файл идентифицированных событий (ФИС). После этого из него удаляются все повторные измерения, строятся распределения исследуемых величин, вычисляются их значения и ошибки, по которым можно судить о надежности и достоверности полученных результатов. И если в полученных данных нет каких-либо систематических искажений и ошибок, то они переписываются на ЛСР.

В ходе анализа результатов измерений камерных фотографий заметная часть событий (10-15%) бракуется в ходе анализа, и после дополнительного просмотра они либо измеряются вторично, либо переводятся в разряд неизмеримых. Поэтому часть уже обрабатывающихся событий снова поступает на вход системы. Для того, чтобы проводить обработку этих событий по уже описанной процедуре без проведения каких-либо ручных операций, для перемеряемых событий составляется специальный

вариант ФРП. В последний из первоначального файла результатов просмотра переписываются лишь те события, информация о которых не переписана на ЛСР, или события не помеченные как неизмеримые. Дальнейший анализ переменных событий ведется точно так же, как и первоначальных измерений, и в принципе этот процесс можно повторять столько раз, сколько потребуется.

Таким образом, в рамках рассмотренной схемы обработки исходные данные для формирования лент суммарных результатов находятся на файлах двух типов - результатов просмотра и результатов реконструкции и идентификации. Это обстоятельство в сочетании с методикой программного отбора событий с учетом результатов визуальных оценок ионизаций треков событий позволило значительно упростить организацию процесса обработки и заменить ряд ручных операций программным решением соответствующих задач. Упрощение процедуры общения пользователя с ЭВМ достигнуто за счет использования небольшого набора управляющих процедур.

§ 3. Основные задачи, решенные в ходе разработки системы

Разработка и реализация рассмотренной в предыдущем параграфе системы потребовали решения на машинной основе ряда задач процесса обработки, среди которых следует отметить следующие:

I. Разработать методику программного отбора физических гипотез с учетом визуальных оценок ионизаций треков, производимых при предварительном просмотре. Для ее реализации потребовалось "научить" ЭВМ отождествлять результаты просмотра и измерений событий и находить соответствие между восстановленными треками и их проекциями на стереоснимках, а также учитывать влияние геометрических факторов на визуальные оценки ионизаций. Внедрение этой методики в практику работы позволило существенно упростить организацию процесса обработки, в несколько раз сократить число вторично просматриваемых событий и исключить из процесса анализа карты отбора гипотез.

II. Разработать достаточно простую технологическую схему процесса обработки единую как для основной массы измеренных событий, так и перемеров.

III. Упрощение процедуры общения пользователей с ЭВМ за счет использования небольшого набора управляющих процедур и расширения возможностей программ системы.

IV. Автоматизировать процедуру формирования ЛСР путем исключения из нее карт отбора и переложения на специально созданные программы таких традиционно человеческих операций, как поиск результатов нескольких измерений одних и тех же событий и выбора из них

наилучших, "просмотра" результатов реконструкции и составления списков событий для повторных измерений, дважды забракованных при реконструкции и идентификации и, наконец, хороших событий, т.е. успешно прошедших все испытания и готовых к переписи на ЛСР. Решение этих задач потребовало разработки соответствующих алгоритмов.

У. Решить задачи систем типа INDEX и STATUS практически без помощи специальных программ и дополнительных файлов^{7,8/}. Действительно, в нашем случае, после завершения обработки для каждой ЛРИ составляется паспорт результатов счета для событий каждой из фото-пленок. Что касается получения данных по всей статистике, то для этой цели используются файлы результатов просмотра. Кроме того, благодаря переходу на унифицированный формат системы "Гидра" и записи головных рекордов событий на файлы результатов даже для забракованных событий, сами результаты счета содержат "историю" обработки каждого поступившего на вход системы события. Таким образом, получение оперативной информации о состоянии дел с обработкой можно получать из имеющихся данных, не прибегая к созданию специализированных программ.

VI. Создание условий для внедрения в практику обработки फिल्मовой информации программного сопровождения этого процесса, реализованного к настоящему времени в основном на этапе измерений, а также при решении ряда других задач в специализированных системах. Суть программного сопровождения в данном случае заключается в том, что с помощью специальных программ производится контроль исходных данных и получаемых результатов на следующих основных этапах процесса обработки:

- измерения;
- реконструкция и идентификация событий;
- программный отбор физических гипотез;
- объединения лент суммарных результатов, полученных в различных лабораториях.

На основании полученных данных программным путем решается вопрос о пригодности данных для последующего анализа. Созданное к настоящему времени программное обеспечение позволяет получать точностные и физические характеристики накапливаемых данных и вырабатывать критерии для организации программного управления процессом прохождения данных. Это достаточно сложный вопрос, который требует проведения специальных исследований для выработки критериев оценки достоверности получаемых результатов и их пригодности для дальнейшего анализа.

УП. Решение всех указанных выше задач позволили существенно повысить надежность системы по отношению к внешним сбоям, под которыми здесь понимаются сбои в работе ЭВМ, ошибки пользователей и т.п.

Заключение

Рассмотренная в работе система математической обработки फिल्मовой информации испытана в ходе анализа данных $d\mu$ -эксперимента. С ее помощью в настоящее время завершается обработка 10 тыс. случаев $d\mu$ -взаимодействий.

Разработанные в ходе ее создания приемы и методы решения задач процесса обработки फिल्मовой информации выходят за рамки конкретного эксперимента и могут быть использованы в других системах обработки экспериментальных данных, в частности, для анализа данных с гибридных систем.

Что касается вопросов дальнейшего развития системы, то здесь необходимо решить задачу организации исходных данных и результатов счета в рамках концепции базы данных. Это позволит автоматизировать процедуру формирования заданий системе и сократит количество информации о текущем состоянии дел, которая в настоящее время фиксируется на бумаге.

Следует указать, что разработка такого типа систем невозможна без тесного сотрудничества прикладных и системных программистов, физиков и инженеров ЭВМ, ибо только такое сотрудничество позволяет решать задачи, возникающие в ходе проведения такого типа разработок.

В заключение авторы выражают глубокую благодарность Н.Ю.Шириковой за чрезвычайно полезные дискуссии по вопросам надежности работы программ и эффективному использованию ЭВМ CDC-6500.

Литература

1. Буздавина Н.А. и др. ОИЯИ, I-81-830, Дубна, 1981.
2. Frenkiel P. et al. В кн.: Methods in subnuclear Physics. v.4, p.3, New-York, 1970.
3. Бадалян С.Г. ОИЯИ, P10-12879, Дубна, 1979.
4. Глаголев В.В. и др. ОИЯИ, P10-82-460, Дубна, 1982.
5. Бадалян С.Г. и др. ОИЯИ, I0-II448, Дубна, 1978.
6. Глаголев В.В. и др. ОИЯИ, P10-82-211, Дубна, 1982.
7. Villemeos P. CERN, N71-6, p.105, Geneva, 1971.
8. Глаголев В.В. и др. ОИЯИ, 81-699, Дубна, 1981.

Рукопись поступила в издательский отдел
5 июля 1985 года.

Принимается подписка на препринты и сообщения Объединенного института ядерных исследований.

Установлена следующая стоимость подписки на 12 месяцев на издания ОИЯИ, включая пересылку, по отдельным тематическим категориям:

ИНДЕКС	ТЕМАТИКА	Цена подписки на год
1.	Экспериментальная физика высоких энергий	10 р. 80 коп.
2.	Теоретическая физика высоких энергий	17 р. 80 коп.
3.	Экспериментальная нейтронная физика	4 р. 80 коп.
4.	Теоретическая физика низких энергий	8 р. 80 коп.
5.	Математика	4 р. 80 коп.
6.	Ядерная спектроскопия и радиохимия	4 р. 80 коп.
7.	Физика тяжелых ионов	2 р. 85 коп.
8.	Криогеника	2 р. 85 коп.
9.	Ускорители	7 р. 80 коп.
10.	Автоматизация обработки экспериментальных данных	7 р. 80 коп.
11.	Вычислительная математика и техника	6 р. 80 коп.
12.	Химия	1 р. 70 коп.
13.	Техника физического эксперимента	8 р. 80 коп.
14.	Исследования твердых тел и жидкостей ядерными методами	1 р. 70 коп.
15.	Экспериментальная физика ядерных реакций при низких энергиях	1 р. 50 коп.
16.	Дозиметрия и физика защиты	1 р. 90 коп.
17.	Теория конденсированного состояния	6 р. 80 коп.
18.	Использование результатов и методов фундаментальных физических исследований в смежных областях науки и техники	2 р. 35 коп.
19.	Биофизика	1 р. 20 коп.

Подписка может быть оформлена с любого месяца текущего года.

По всем вопросам оформления подписки следует обращаться в издательский отдел ОИЯИ по адресу: 101000 Москва, Главпочтамт, п/я 79.

Балгансурен Я. и др.

Высокоавтоматизированная система математической обработки फिल्मовой информации

P10-85-516

В ходе проводимых в ОИЯИ работ по повышению уровня автоматизации процесса обработки फिल्मовой информации была разработана устойчивая к внешним сбоям система математической обработки экспериментальных данных, обеспечивающая автоматический отчет данных по заданным цепочкам программ, разбор и предварительный анализ результатов счета. Обработка данных $\alpha\beta$ -эксперимента с помощью этой системы показала, что она обеспечивает стабильную и практически безаварийную обработку данных, резкое сокращение затрат ручного труда на разбор и анализ результатов счета. "Общение" пользователей с программами системы производится с помощью ограниченного набора управляющих процедур.

Работа выполнена в Лаборатории вычислительной техники и автоматизации ОИЯИ.

Сообщение Объединенного института ядерных исследований. Дубна 1985

Перевод О.С.Виноградовой

Balgansuren J. et al
High-Automated System for Film Data Mathematical Processing

P10-85-516

The mathematical film data processing system has been developed. It is failproof and provides a possibility to process data automatically by the chains of programs and to analyse preliminarily computer results. The processing of $\alpha\beta$ -experiment data using this system has shown that it provided stable and practically nonaccidental data processing, and significantly reduced manual work expenses for analysing computer results. The user's "contact" with system programs goes via a limited set of controlled procedures.

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

Communication of the Joint Institute for Nuclear Research. Dubna 1985