

УДК 681.3.019; 681.786.42

## СИСТЕМА ОЦИФРОВКИ ГРАФИКОВ (СОГ)

*А.П.Иерусалимов, Е.А.Иерусалимов\*,  
Ж.Ж.Мусульманбеков, Э.Г.Никонов, В.С.Рихвицкий*

Представлена система оцифровки графиков (СОГ), включающая аппаратную и программную части. Рассматриваются два варианта ее реализации: с использованием полуавтоматического измерительного устройства (ПУОС) и с использованием сканера. Приводятся данные оцифровки тестовых графиков и анализируются полученные результаты.

Работа выполнена в Лаборатории вычислительной техники и автоматизации ОИЯИ.

## System of Graphics Digitization (SGD)

*A.P.Ierusalimov et al.*

The system of graphics digitization (SGD) consisting of hard- and software is presented in this paper. Two variances of its realization are described: using semiautomatic device and using scanner. The data of digitization of test graphics and obtained results are analysed.

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

### 1. Введение

В настоящее время опубликованные экспериментальные данные, представляющие интерес для научного сообщества в физике низких, промежуточных и высоких энергий, заносятся в виде таблиц в соответствующие базы данных. В физике низких энергий — это банки ядерных данных, поддерживаемые Национальным ядерным центром BNL (NNDC) и Европейским агентством по ядерной энергии (NEA). В физике промежуточных и высоких энергий измеренные данные заносятся в базу REACTION системы баз данных по физике частиц (PPDS, HEPDATA), поддерживаемых сотрудничеством институтов в составе ИФВЭ, ИТЭФ, ОИЯИ, RAL (Durham) и Particle Data Group (Berkeley) [1].

Используя подобные фактографические базы, физики получают возможность автоматизировать процедуру выборки числовых данных для их последующего применения, а именно для сравнения с другими данными или модельными расчетами и т.п.

Однако с самого начала эксплуатации фактографических баз данных существует проблема, не решенная до настоящего времени. Обычно измеренные и опубликованные в виде графиков данные предоставляются по запросу авторами публикаций для внесения в соответствующую базу в оцифрованном виде. При этом около 10 % запро-

---

\*МИФИ, Москва

сов сотрудников, обслуживающих эти базы данных, не удовлетворяются, что связано с занятостью авторов, потерей информации, отсутствием промежуточных таблиц и т.п. Это значительно снижает ценность такого способа хранения и доступа к фактографической информации и, соответственно, интерес пользователя к таким системам, поскольку в таком случае он в ответ на свою поисковую формулу получает пустой файл.

В результате физик вынужден, используя подручные средства (линейку и пр.), кропотливо переносить точку за точкой на миллиметровку, а затем восстанавливать с нее табличные значения. Это может привести как к порче журнальных страниц, так и к значительной потере времени, не гарантируя при этом необходимой точности восстановления точек.

Поэтому становится актуальной проблема оцифровки графической информации (графиков, гистограмм и пр.), опубликованной в журнальных статьях и препринтах и не имеющей табличных эквивалентов.

В настоящее время реализованы два варианта решения задачи оцифровки графических данных:

- 1) использование приборов ПУОС (полуавтоматическое устройство обмера снимков);
- 2) использование сканеров достаточно высокого разрешения (не менее 600 dpi) с последующим восстановлением цифровой информации с помощью разработанного для этого математического обеспечения на PC в среде WINDOWS.

## 2. Оцифровка на ПУОС

Измерительная система ПУОС [2] предназначена для обмера фотоснимков с трековых детекторов. Максимальный размер обрабатываемого изображения  $\sim 200 \times 200$  мм. С помощью оптической системы увеличенный участок изображения проецируется на матовый экран диаметром 170 мм. Коэффициент увеличения изображения можно варьировать от 10 до 50 с помощью сменных оптических объектов, что позволяет надежно измерять изображения с очень мелким масштабом.

Оператор-измеритель производит рекогносцировку изображения, его перемещение и точное наведение измерительной марки с помощью измерительного стола. Для отсчета каждой из координат используются бесконтактные датчики на дифракционных решетках. Измерительная система ПУОС характеризуется высокой точностью (цена отсчета по каждой из координат — 2,5 мкм) и линейностью.

Прием цифровой информации и некоторых служебных меток осуществляет PC с помощью специальной электронной платы. За 4-часовую смену оператор может измерить  $\sim 1500$  точек.

Процедура оцифровки включает следующие этапы:

- 1) ксерокопирование рисунка;
- 2) измерение оператором ПУОСа «полезных» точек на копии;
- 3) программная обработка измеренных точек (создание числового эквивалента измеренного графика или гистограммы).

Предварительные тестовые оценки по оцифровке типичных графиков, взятых из журнальных статей, показали, что погрешность восстановления точек для графиков размером  $60 \times 60$  мм составляет  $\sim 10$  мкм.

Для реализации этого варианта необходимо иметь один или несколько приборов ПУОС на линии с РС типа Pentium и ксерокопировальный аппарат. Имеющееся математическое обеспечение ПУОСа потребует незначительной модификации.

### 3. Оцифровка в системе сканер — РС

Процедура оцифровки состоит из следующих шагов:

- 1) копирование рисунка с помощью сканера для записи графических данных в память РС или на внешний магнитный носитель;
- 2) предварительная обработка графических данных с помощью развитого графического редактора;
- 3) измерение оператором с помощью «мыши» отредактированного графического рисунка, выведенного на экран монитора РС;
- 4) программная обработка измеренных точек (создание табличного эквивалента измеренного графика или гистограммы).

Если исходный рисунок представлен в виде PostScript- или PDF-файла, то необходимость в первом шаге процедуры оцифровки отпадает. Такое представление рисунков, типичное для современных публикаций, является предпочтительным, поскольку потери точности, обусловленные процессом подготовки публикации и последующим копированием рисунка при сканировании, исключаются.

Для реализации этого варианта необходим РС типа Pentium, сканер высокого разрешения (не менее 600 dpi) и соответствующее МО процесса оцифровки. Как в случае применения прибора ПУОС, оператор может оцифровать ~1500 точек графического изображения за 4 часа работы.

Кроме оцифровки графиков непосредственно оператором-измерителем, для этого варианта реализован режим автоматической оцифровки непрерывных кривых. В этом режиме оператору достаточно отметить только начальную и конечную точки кривой. Специальная программа осуществляет все необходимые операции.

### 4. Предварительные результаты оцифровки

Для обоих вариантов была проведена пробная оцифровка графических данных. График, взятый из журнальной статьи [3], был ксерокопирован на прозрачную пленку, измерен на приборе ПУОС, а затем полученные данные обрабатывались на ЭВМ с помощью соответствующей программы. На рис.1 показан график, предоставленный авторами статьи (слева), и график, построенный по результатам оцифровки (справа). В табл.1 приведены исходные (оригинальные) и оцифрованные (восстановленные) данные. Из сравнения рисунков и анализа таблицы можно сделать вывод о хорошем согласии между исходными и восстановленными данными.

Другой график из той же статьи был просканирован с помощью сканера HP SCANJET 3С и обработан по специальной программе оцифровки на персональной ЭВМ. Полученные результаты представлены на рис.2 и в табл.2. В этом случае также наблюдается хорошее согласие между исходными и восстановленными данными.

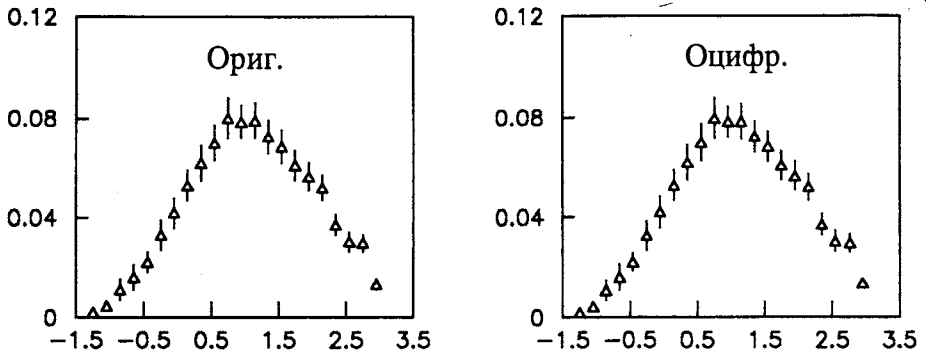


Рис.1

Таблица 1

$N_T$	$X_{\text{ориг.}}$	$X_{\text{восст.}}$	$Y_{\text{ориг.}}$	$Y_{\text{восст.}}$	$\sigma Y_{\text{ориг.}}$	$\sigma Y_{\text{восст.}}$
1	-1,25	-1,248	0,002	0,0016	0,001	0,0015
2	-1,05	-1,049	0,005	0,0046	0,002	0,0022
3	-0,85	-0,858	0,011	0,0105	0,004	0,0039
4	-0,65	-0,653	0,016	0,0157	0,005	0,0049
5	-0,45	-0,456	0,022	0,0218	0,004	0,0035
6	-0,25	-0,254	0,033	0,0324	0,006	0,0058
7	-0,05	-0,052	0,042	0,0421	0,006	0,0064
8	0,15	0,151	0,053	0,0527	0,006	0,0059
9	0,35	0,354	0,062	0,0619	0,007	0,0070
10	0,55	0,553	0,070	0,0700	0,007	0,0073
11	0,75	0,752	0,080	0,0797	0,008	0,0080
12	0,95	0,957	0,079	0,0788	0,007	0,0066
13	1,15	1,152	0,079	0,0786	0,007	0,0069
14	1,35	1,348	0,073	0,0730	0,007	0,0065
15	1,55	1,549	0,069	0,0691	0,007	0,0065
16	1,75	1,747	0,061	0,0609	0,006	0,0061
17	1,95	1,949	0,057	0,0571	0,006	0,0063
18	2,15	2,147	0,052	0,0522	0,005	0,0052
19	2,35	2,348	0,037	0,0369	0,004	0,0037
20	2,55	2,547	0,030	0,0301	0,004	0,0038
21	2,75	2,762	0,029	0,0288	0,004	0,0040
22	2,95	2,952	0,013	0,0132	0,002	0,0018

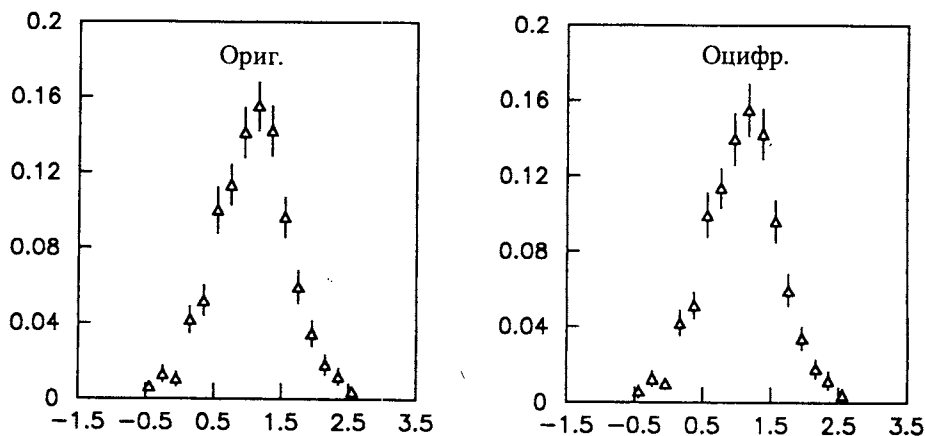


Рис.2

Таблица 2

$N_T$	$X_{\text{ориг.}}$	$X_{\text{восст.}}$	$Y_{\text{ориг.}}$	$Y_{\text{восст.}}$	$\sigma Y_{\text{ориг.}}$	$\sigma Y_{\text{восст.}}$
1	-0,45	-0,436	0,006	0,0053	0,003	0,0036
2	-0,25	-0,238	0,013	0,0125	0,004	0,0039
3	-0,05	-0,042	0,011	0,0105	0,004	0,0033
4	0,15	0,165	0,042	0,0420	0,007	0,0065
5	0,35	0,372	0,052	0,0515	0,008	0,0069
6	0,55	0,567	0,100	0,1005	0,012	0,0117
7	0,75	0,763	0,114	0,1144	0,011	0,0109
8	0,95	0,958	0,141	0,1397	0,013	0,0134
9	1,15	1,167	0,156	0,1556	0,013	0,0135
10	1,35	1,374	0,142	0,1423	0,013	0,0129
11	1,55	1,565	0,097	0,0967	0,011	0,0114
12	1,75	1,757	0,059	0,0591	0,008	0,0078
13	1,95	1,954	0,035	0,0344	0,007	0,0066
14	2,15	2,162	0,018	0,0178	0,005	0,0047
15	2,35	2,345	0,012	0,0117	0,004	0,0045
16	2,55	2,551	0,004	0,0040	0,002	0,0029

На рис.3 показан результат автоматической оцифровки кривой из статьи [4]. Звездочками отмечена каждая десятая восстановленная точка.

Предварительное сравнение двух методов оцифровки позволяет сделать следующие выводы:

1. На приборах ПУОС возможна с достаточной точностью оцифровка графиков практически любой степени сложности.

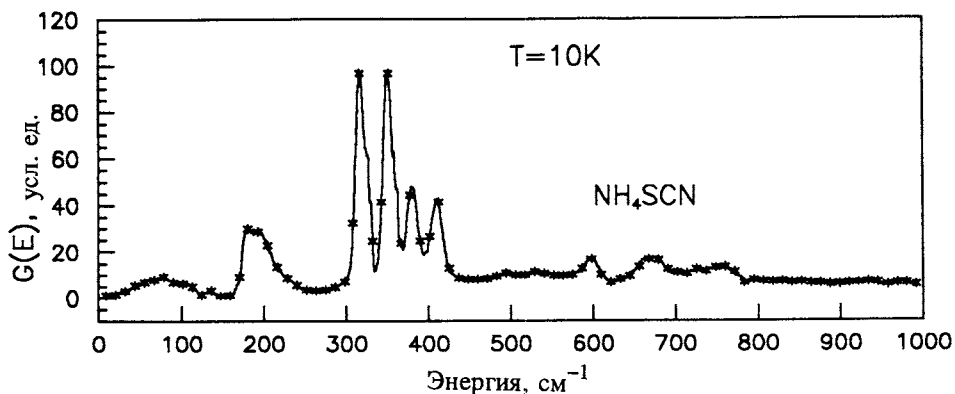


Рис.3

2. Более жесткие требования предъявляются к графикам при оцифровке с помощью сканера: достаточный размер изображения (не менее  $30 \times 30$  мм), хорошая контрастность, неперегруженность мелкими деталями.

3. Из расчета ежедневной 4-часовой работы оператора-измерителя можно добиться как в первом, так и во втором вариантах оцифровки не менее 5 тыс. графиков в год на одного оператора (исходя из 50 точек на график).

## 5. Заключение

Таким образом, предлагаемая система оцифровки графиков (СОГ) позволит ликвидировать существующую неполноту мировых фактографических баз данных по физике частиц. Она поможет освободить исследователей от непроизводительной и трудоемкой ручной работы, связанной с оцифровкой необходимых графиков и гистограмм, не внесенных по каким-либо причинам в базы данных.

Каждый из предложенных вариантов оцифровки, имея определенные достоинства и недостатки, взаимно дополняет друг друга. Система оцифровки с ПУОСом, обладающая высокой точностью измерения и восстановления точек, что позволяет ее использовать для обработки сложных рисунков, является в то же время громоздкой и нетиражируемой. Второй вариант оцифровки, будучи менее точным, легко тиражируется и позволяет использовать стандартное оборудование при наличии соответствующего программного обеспечения.

## Литература

1. Review of Particle Physics. Phys. Rev., D54, 1996.
2. Виноградов А.Ф. и др. — ОИЯИ, 10-8783, Дубна, 1975.
3. Бекмирзаев Р.Н. и др. — ЯФ, 1995, т.58, с.63.
4. Длоуга М. и др. — ОИЯИ, P14-96-157, Дубна, 1996, с.5.

Рукопись поступила 19 августа 1997 года.