

Л и т е р а т у р а

1. Т и х о н о в А.Н., А р с е н и н В.Я. Методы решения некорректных задач. М., "Наука", 1974.
2. П ш е н и ч н и к о в В.В. Исследования процедуры восстановления в частотной области.
В сб.: Автоматизация экспериментальных исследований. Куйбышев, КуАИ, 1975.
3. Г о л д Б., Р е й д е р Ч. Цифровая обработка сигналов. М., "Советское радио", 1973.

Е.Ю. Ларцев

АНАЛИЗ КАЧЕСТВА ДАННЫХ В АВТОМАТИЗИРОВАННОЙ СИСТЕМЕ ОБРАБОТКИ ИЗМЕРИТЕЛЬНОЙ ИНФОРМАЦИИ

Рост сложности экспериментальных исследований как в научной, так и аппаратурной части вызывает появление огромных потоков измерительных данных, предварительную обработку которых зачастую необходимо проводить в сравнительно сжатые сроки. Это определяет необходимость применения систем автоматизации процесса обработки, базирующихся на мощных УЦВМ. Целью предварительной обработки является исключение специфического воздействия систем сбора и передачи данных на измерительную информацию. При этом необходимой составной частью почти каждой фазы предварительной обработки (измерительные данные проходят до IO фаз обработки) является анализ качества данных, алгоритмы которого строятся с учетом статистических характеристик обрабатываемых данных и априорных сведений о форме их представления. На основании результатов этого анализа на конкретной фазе принимается решение о проведении последующих фаз обработки и очередности их выполнения [1].

В данной работе предлагается программный комплекс, реализующий процедуру анализа качества данных на фазе локализации импульсных помех (сбоев), применительно к системе автоматизированной обработки многоканальной измерительной информации.

Рассматриваемый комплекс программ предназначен для выделения из обрабатываемых данных импульсных и флуктуационных помех с целью анализа их статистических характеристик, на основании которых производится эффективный выбор параметров алгоритмов фильтрации импульсных помех.

Для удобства изложения вопроса предложим следующую классификацию измерительных сигналов и сопутствующих им помех. Составляющие интегральный поток обрабатываемых данных измерительные сигналы по характеру их изменения можно разбить на два класса:

1. Функционально-переменные сигналы, имеющие некоторое конечное приращение сигнала.

2. Функционально-постоянные сигналы с пренебрежимо малыми приращениями (к ним относятся сигналы типа уровневых и калибровочных).

Различного рода шумы и помехи присутствующие в реальном сигнале также можно объединить в две группы помех:

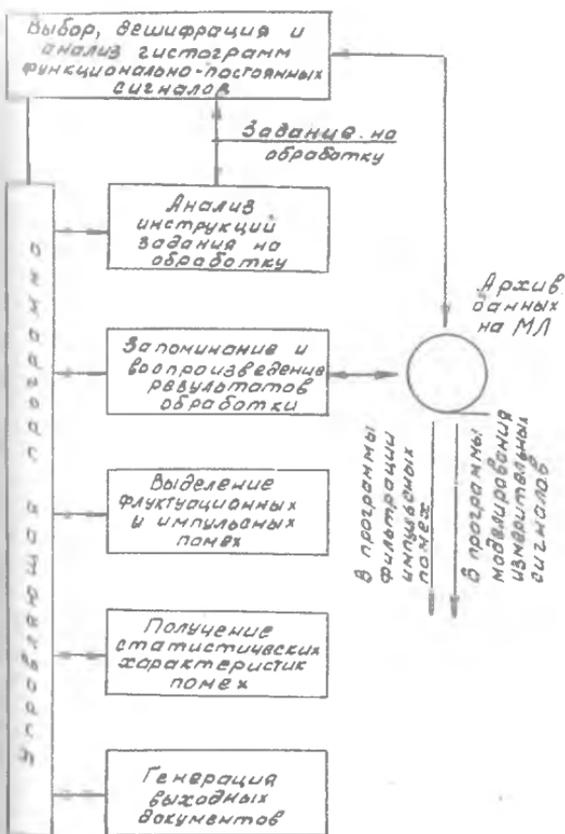
1. Импульсные помехи, характеризующиеся большой амплитудой и сравнительно малой частотой своего появления.

2. Флуктуационные помехи, отличающиеся малой амплитудой и присутствующие практически в каждой выборке сигнала.

Будем считать импульсными помехи, амплитуда которых превышает $3\sigma_{\varphi}$, где σ_{φ} - среднеквадратичное отклонение флуктуационной помехи φ .

Предлагается следующий подход к определению статистических характеристик помех. Весь поток измерительных сигналов разбивается на группы с идентичными условиями сбора данных и передачи их по каналу связи. С большой степенью достоверности можно допустить что статистические характеристики помех, присутствующих в каждом из сигналов группы, будут близки между собой. Поэтому для получения указанной статистической информации предлагается использовать только функционально-постоянные сигналы, при обработке которых процедура выделения импульсных и флуктуационных помех наиболее проста. Получаемые при этом результаты распространяются и на функционально-изменяющиеся сигналы данной группы.

На рис. 1 представлена структурная схема анализатора качества данных (АКД) - программного комплекса, реализующего предложен-



Р и с. 1. Структурная схема АКД

2-й этап. На данном этапе производится амплитудное разграничение импульсной и флюктуационной помех. При анализе полученной ранее гистограммы значений сигнала вычисляются оценки среднего значения сигнала \hat{M}_s и среднеквадратичного отклонения флюктуационной помехи $\hat{\sigma}_f$. Далее организуется итерационная процедура поочередного уточнения \hat{M}_s и $\hat{\sigma}_f$. При этом из рассмотрения исключается часть гистограммы, соответствующая значениям сигнала, искаженного импульсной помехой, т.е. значениям, лежащим вне интервала $[\hat{M}_s - 3\hat{\sigma}_f, \hat{M}_s + 3\hat{\sigma}_f]$. Итерационная процедура продолжается до момента стабилизации значений \hat{M}_s и $\hat{\sigma}_f$. После чего \hat{M}_s и $\hat{\sigma}_f$ считаются определенными и, как следствие, считаются определенными условные амплитудные границы между импульсной и

ный подход к получению статистической информации о помехах. Обработка данных в АКД проводится в несколько этапов

1-й этап. Из совокупности измерительных сигналов выделяются сигналы функционально-постоянного типа, производится их дешифрация и представление в виде, удобном для дальнейшей обработки (вследствие необходимости работы с битовой структурой данных программы, реализующие данный этап написаны на АВТОКОДе, остальные же программы АКД - на ФОРТРАНе). Одновременно производится построение гистограмм значений выделенных сигналов.

флуктуационной помехами.

3-й этап. Определяются статистические характеристики помех, знание которых необходимо для повышения эффективности процедуры фильтрации импульсных помех:

$W(A_n)$ - гистограмма распределения плотности вероятности амплитуд импульсных помех, начиная с $3\sigma_{\xi}$;

$W(A_{\xi})$ - гистограмма распределения плотности вероятности амплитуд флуктуационных помех;

$W(N_g)$ - гистограмма распределения плотности вероятности количества достоверных выборок сигнала между импульсными помехами;

$W(N_n)$ - гистограмма распределения плотности вероятности размера пакета импульсных помех, а также два первые моменты и граничные значения каждого из распределений.

По окончании 3-го этапа результаты анализа с помощью соответствующих программ выдаются на печать и записываются в архив данных на магнитной ленте (там же хранятся дешифрованные на I-ом этапе обработки последовательности отсчетов обрабатываемых сигналов и промежуточные результаты анализа, необходимые для его продолжения в случае прерывания процесса обработки). В дальнейшем статистическая информация, полученная АКД, используется программами фильтрации импульсных помех и программами моделирования измерительных сигналов.

Программный комплекс АКД реализован на языках АВТОКОД и ФОРТРАН (транслятор ФМ-20) в рамках операционной системы ОСПО ЭВМ М-222. В теках системной ленты созданы библиотеки текстов программ и модулей загрузки, каждая из которых содержит 19 разделов (по числу программ, входящих в АКД). Активизация модулей загрузки или текстовых программ осуществляется в соответствии с директивами языка управления заданиями управляющей программой монитор ОСПО, а конфигурация АКД определяется его управляющей программой на основании инструкций задания на обработку.

Л и т е р а т у р а

1. Евдокимов В.П., Покрасс В.М.
Методы обработки данных в научных космических
экспериментах. М., "Наука", 1977.

И.С. Ризаев

СЖАТИЕ БАЗЫ ДАННЫХ

Запоминающие устройства с прямой записью (ЗУПД) позволяют создавать информационные системы, базирующиеся на единой интегрированной базе данных. Для эффективной работы с базой данных необходимо, чтобы прямой доступ был обеспечен в любое время ко всей базе. Но ввиду ограниченности дисковой памяти и больших объемов информации такой доступ, как правило, не обеспечивается. Решением этой проблемы может быть сокращение избыточности хранимой информации в базе данных путем уплотнения или сжатия этих данных.

Пусть $M = \{m_i\}$ - множество массивов информации, составляющих базу данных до сжатия; $M' = \{m'_i\}$ - множество полученных массивов информации после сжатия данных; $A = \{a_i\}$ - алгоритмы преобразования, $i = \overline{1, n}$. Видимо, сжатие данных можно проводить только в том случае, если будет выполняться соотношение

$$M' + A < \bar{M},$$

где \bar{M} , M' , A - объемы соответствующих массивов и алгоритмов.

Таким образом, задача сводится к созданию или подбору соответствующих алгоритмов сжатия. Существуют различные методы сжатия: машинные, немашинные, методы оптимального размещения, комбинированные. Хотя методы сжатия разрабатываются давно, применительно к базам данных они стали рассматриваться сравнительно недавно [1]. Большинство известных методов сводятся к рассмотрению сжатия информации, расположенной непосредственно в памяти ЭВМ. Чаще используют лексическое кодирование [2], суть которого заключается в том, что более длинному сообщению ставится в соответствие более короткое, а именно, порядковый номер данного сообщения. В справочниках или в памяти ЭВМ хранятся два словаря: