

В. С. КИРИЧУК, Б. Н. ЛУЦЕНКО,  
Б. М. ПУШНОЙ, Г. П. ЧЕЙДО  
(Новосибирск)

## ОБ АЛГОРИТМАХ АВТОМАТИЧЕСКОЙ ОБРАБОТКИ ИНТЕРФЕРОГРАММ

Интерферограмму можно рассматривать как один из элементов некоторой измерительной системы, в которую входят источник излучения, устройства, формирующие световые потоки, фотопластинка, регистрирующая интерференционную картину, сканирующая система, электронная вычислительная машина. Способ действия этой системы определяется в конечном счете алгоритмом, который не только задает характер и порядок математических операций с массивом данных, но и режим сканирования [1—3]. Ниже рассматриваются основные требования к алгоритмам обработки интерферометрической информации в автоматических системах и делается попытка определить круг специальных задач, связанных с созданием таких алгоритмов.

Введем следующие обозначения: рабочая волна интерферометра  $\lambda = 2\pi/\omega$ ; разность хода  $Z(x, y)$ ; координаты точки интерферограммы  $x, y$ ; освещенность фотослоя при регистрации интерферограммы  $I(x, y)$ ; плотность почернения (сигнал на выходе системы сканирования)  $D(x, y)$ ; математическое ожидание  $D(x, y)$  по ансамблю  $M(x, y)$ ; характеристическая кривая (математическое ожидание сигнала плотности как функция освещенности)  $M_1 = F_1(I)$ ; характеристическая кривая дисперсии результирующего шума гранулярности и тракта сканирования  $M_2 = F_2(I)$ ; характеристические кривые высших моментов распределения шума  $M_k = F_k(I)$ .

Представляется удобным рассматривать следующую математическую модель интерферограммы:

$$D(x, y) = M(x, y) + \xi(x, y),$$

где

$$M(x, y) = F_1[I(x, y)];$$

$$I(x, y) = A(x, y) + B(x, y) \cos \omega Z(x, y);$$

$\xi(x, y)$  — случайная функция с нулевым средним ( $M_1 = 0$ ), дисперсией  $M_2 = F_2[I(x, y)]$  и высшими моментами  $M_k = F_k[I(x, y)]$ .  $A(x, y)$  и  $B(x, y)$  выражают неравномерность интенсивности опорного и информационного пучков, с которыми в ряде случаев приходится считаться.

Общей целью обработки интерферограмм является получение характеристик некоторой функции  $Z(x, y)$ , описывающей исследуемое явление. В большинстве случаев эти характеристики связаны со способом математического описания, принятым в теоретических исследованиях. Форма окончательного результата и характер преобразований зафиксированной на интерферограмме информации чрезвычайно разнообразны и часто изменяются при проведении исследований одного и того же объекта. Однако при всем своем разнообразии преобразования содержат общую операцию, сводящуюся к получению достаточно хороших приближений  $\hat{Z}(x, y)$ , описывающих полный фазовый набег  $Z(x, y)$  в пределах всей интерферограммы или какой-то наиболее содержательной ее области. Поэтому общей формой обработки интерферограмм

следует считать оценку аргумента зарегистрированной на ней периодической функции.

Все специфические трудности организации автоматической обработки интерферограмм обусловлены тем, что интерферографическая измерительная система имеет периодическую функцию преобразования [4]. Как известно, любой измерительный преобразователь может быть охарактеризован функцией, связывающей измеряемую величину  $Z$  с регистрируемой  $D$ . Эта функция определяется физическим явлением, положенным в основу работы преобразователя. Например, тепловым преобразователям соответствует квадратичная функция преобразования. У интерферометрических преобразователей эта функция периодическая (рис. 1). Измеряемой величиной является смещение фазы информационного пучка относительно опорного в некоторой точке интерферограммы. Регистрируется на фотоносителе некоторая функция результирующей амплитуды колебаний. Длина рабочей волны интерферограммы играет роль масштаба преобразования.

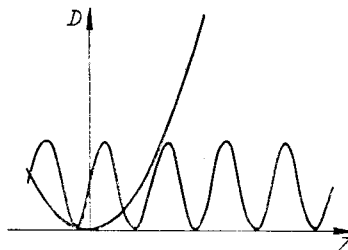


Рис. 1.

Таким образом, если известна характеристическая кривая фотоматериала, нетрудно установить связь между плотностью изображения в данной точке и разностью фаз опорного и информационного пучков. Однако зависимость  $\hat{Z}(x, y)$  от  $M_1(x, y)$  является неоднозначной. Разность фаз определяет  $\hat{Z}(x, y)$  с точностью до целого числа фазовых циклов, и, кроме того, как нетрудно заметить, одному и тому же уровню плотности соответствуют два значения фазы:  $\varphi$  и  $\pi - \varphi$ .

Обычно интервал изменения измеряемой величины в пределах интерферограммы достаточно широк, и на ней регистрируется большое количество циклов. В задачу обработки входит определение полной разности фаз с учетом целого числа циклов, чтобы функция  $\hat{Z}(x, y)$  в пределах интерферограммы не содержала разрывов и разрывов первой производной, обусловленных неоднозначностью.

Задачу обработки интерферограмм будем рассматривать применительно к наиболее сложному случаю, который характеризуется следующими особенностями:

а) в пределах одной интерферограммы длина периода может изменяться в широких пределах, так что минимальный период близок к границе разрешающей способности системы сканирования, а наибольшее значение периода может быть соизмеримо с размерами интерферограммы;

б) функция  $Z(x, y)$  или ее производные могут иметь разрывы;

в) характеристическая функция фотоматериала априори не задана и подлежит определению в ходе обработки интерферограммы;

г) шум интерферограммы может быть значительным, дисперсия может достигать значений  $1/12$  от квадрата рабочего интервала плотности. Такой уровень шума типичен для голографических интерферограмм, используемых при анализе деформаций.

В этих условиях решение задачи автоматизации обработки интерферограмм связано с разработкой алгоритмов, обладающих высокой помехоустойчивостью. Такие алгоритмы должны базироваться на статистических методах оценивания параметров функции  $\hat{Z}(x, y)$ , причем оценивание должно производиться по выборке достаточно большого объема. Как показывает практика, выборка может занимать на интер-

ферограмме такую площадь, на которой располагается несколько циклов зарегистрированной функции.

Для получения статистических оценок функции  $\hat{Z}(x, y)$  ее выражают через совокупность параметров, аппроксимируя ее тем или иным способом. При этом возникают обычные вопросы о выборе вида описания. Однако даже при линейном относительно параметров описании зависимость от измеренных величин оказывается нелинейной, так как мы ищем приближение к аргументу некоторой периодической функции, в простейшем случае — синуса.

Статистические методы, основанные на квадратичном критерии, приводят к необходимости минимизации функционала вида

$$\Phi = \sum_x \left\{ D(x) - F_1 \left[ \hat{A}(x) + \hat{B}(x) \cos \frac{2\pi}{\lambda} \hat{Z}(x) \right] \right\}^2$$

который существенно нелинеен относительно неизвестных параметров. Общие методы решения таких задач основаны на применении итерационных методов, хотя некоторыми приемами удается преобразовать массивы данных и получить статистики, линейные относительно определяемых параметров. Один из таких приемов мы продемонстрируем ниже.

Применение итерационных методов осложняется многоэкстремальностью функционала, наличием глубоких локальных минимумов и зависимостью функционала от количества периодов на интервале обработки.

Проиллюстрировать характер зависимости функционала от параметров можно на примере функционала

$$\sum_{i=1}^{50} [\sin \omega b_1 x_i - \sin \omega (a_0 + a_1 x_i + a_2 x_i^2 + a_3 x_i^3)],$$

фиксируя значения трех параметров  $a_i$  и варьируя четвертым.

На рис. 2 представлены четыре кривые, полученные при  $\Delta x = x_{i+1} - x_i = 0,08$ ;  $\omega = 5$ ;  $b_1 = 2$ . Первая кривая построена при  $a_1 = 2$ ,  $a_2 = a_3 =$

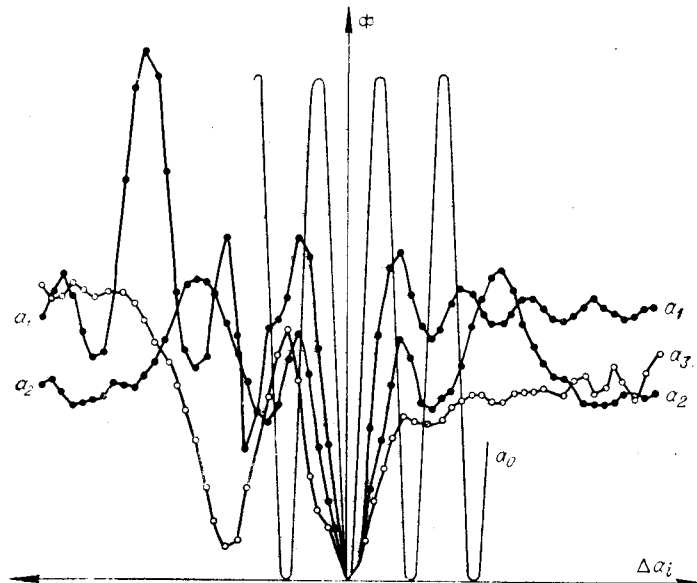


Рис. 2.

$=0$  как функция  $a_0$ . Она может быть выражена как  $1 + \cos \omega a_0$ . По параметру  $a_0$  функционал имеет счетное множество минимумов, равных нулю. Выход на любой из них приводит к решениям, отличающимся одно от другого на целое число периодов  $\tau = 2\pi/\omega$ . В этом проявляется неоднозначность измерительной системы. По параметру  $a_1$  функционал имеет один главный минимум и счетное множество локальных. Если область значений  $x$  достаточно велика по сравнению с периодом  $\tau$  и параметры  $a_0, a_2, a_3$  равны нулю, кривая близка к  $1 - \frac{\sin \omega a_1}{\omega a_1}$ . Последнее выраже-

ние позволяет судить о соотношении величин главного и локальных минимумов функционала при вариации параметра  $a_1$ . Однако при малом интервале анализа, как видно из рис. 2, эта кривая значительно отличается от приведенного выше вида, особенно при отрицательных приращениях  $a_1$ , что соответствует увеличению периода приближающей кривой. Сечение функционала по параметрам  $a_2$  и  $a_3$  имеет в общем такой же характер. Обращает на себя внимание наличие глубокого локального минимума в отрицательной области параметра  $a_3$ . Совершенно очевидно, что при наличии даже сравнительно небольшого шума такой минимум может быть отождествлен с главным, что приведет к существенной ошибке в оцениваемых параметрах. Однако если увеличить интервал анализа хотя бы вдвое, этот опасный локальный минимум практически исчезает, что показано на рис. 3, где приведено сечение того же функционала (кривая  $a$ ) в области отрицательных вариаций  $a_3$  с увеличенным вдвое интервалом анализа (значение  $i$  изменялось от 1 до 100). Отсюда следует, что увеличение интервала анализа, т. е. расширение поля сканирования на возможно большее количество циклов зарегистрированной на интерферограмме функции, приводит не только к повышению точности оценивания параметров за счет увеличения количества обрабатываемых отсчетов, но и к увеличению достоверности получаемых результатов за счет снижения вероятности выхода на локальный минимум функционала.

Форма функционала достаточно сложна, особенно если принять во внимание кривую  $b$  на рис. 3, которая соответствует вариации параметра  $a_3$  при других значениях остальных параметров:  $a_0 = a_2 = 1; a_1 = 2$ . Кривая  $b$  существенно отличается от соответствующей кривой на рис. 2. В связи с этим не представляется возможным использовать какие-либо особенности функционала как признаки близости к главному минимуму. Единственное, что можно получить в данных условиях — указать нижнюю границу значений локальных минимумов (при большом интервале) и оценить аналитически поведение функционала в окрестностях главного минимума, в частности — его ширину. Эти данные необходимы для вычисления дисперсий оценок параметров, размеров области сходимости и оценки сложности задачи поиска главного экстремума. Форма главного минимума определяется, естественно, видом принятого аналитического описания функции  $Z$ . Для сужения области поиска на первом этапе привлекается вся доступная априорная информация о параметрах, включая и визуальный просмотр интерферограммы. Дискретность перебора значений параметра выбирается согласно раскрытию главного минимума и уровню шума. На следующих этапах обработки в роли нулевого

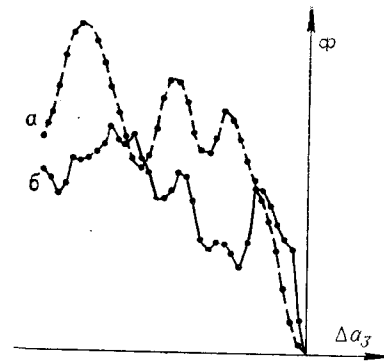


Рис. 3.

приближения могут уже выступать значения параметров, полученные на предыдущем шаге. Разумеется, сдвиг вводимых в обработку массивов данных не должен приводить к выходу параметров из зоны сходимости главного минимума. Алгоритм, использующий такой подход, предусматривает построение достаточно детальной математической модели интерферограммы, содержащей наряду с непосредственно интересующими нас параметрами функции  $Z(x, y)$  и ряд сопутствующих им параметров, описывающих в основном качество используемого тракта преобразования и регистрации информации, как-то: коэффициенты характеристической функции фотоматериала, уровни информационного и опорного потоков. Правда, в случае постоянства сопутствующих параметров по полю интерферограммы информация о них будет поступать с каждого вводимого в обработку фрагмента. Это открывает возможность последующего уточнения их оценок, что в конечном счете дает возможность уточнить и оценки интересующих исследователя параметров. В указанных потенциальных резервах достоинство такого подхода. Недостатки явственны из предшествующих рассуждений: громоздкость вычислений, необходимость задания довольно точного нулевого приближения и опасность выхода на локальный минимум. Разумеется, представляют определенный практический интерес беспойсковые алгоритмы оценивания параметров первого приближения. Однако анализ их эффективности в конечном счете связан со свойствами описанного выше функционала и представляет собой самостоятельную задачу.

Попытаемся найти минимальную величину дисперсии оцениваемых параметров аргумента периодической функции, используя теорему Рао — Крамера — Фреше [5, 6]. В этой теореме утверждается, что корреляционная матрица оценок  $B$  и информационная матрица Фишера  $I$  находятся в следующем соотношении:

$$Z^T B^{-1} Z \leq Z^T / Z, \quad (1)$$

где  $Z$  — текущий вектор, размерность которого равна числу оцениваемых параметров. Неравенство (1) означает, что эллипсоид рассеивания вектора оценок лежит вне эллипсоида, определенного информационной матрицей Фишера. В свою очередь, матрица Фишера выражается через вторую производную функции правдоподобия  $L$  по оцениваемым параметрам

$$I = \left\| E \left( - \frac{\partial^2 \ln L}{\partial a_i \partial a_j} \right) \right\|.$$

Существенное предположение теоремы состоит в том, чтобы область выборочного пространства, в которой плотность распределения отлична от нуля, не зависела от вектора параметров. Применительно к рассматриваемой задаче это означает отсутствие насыщения фотоматериала на интерферограмме и отсутствие ярко выраженной зависимости между шумом негатива и полезным сигналом. Этим требованиям удовлетворяет хорошо «проработанная» интерферограмма, не выходящая за пределы линейного участка характеристической кривой фотоматериала.

Рассмотрим упрощенную модель интерферограммы:

$$D(x) = A + B \cos \left[ \omega \sum_{i=1}^m a_i P_i(x) \right] + \xi(x), \quad (2)$$

где  $A$  и  $B$  — известные постоянные;  $P_i(x)$ ,  $i = \overline{1, m}$  — система линейно независимых функций;  $\xi(x) \in N(0, \sigma^2)$ .

Логарифм функции правдоподобия будет иметь вид

$$L = \text{const} - \frac{1}{2\sigma^2} \sum_{j=1}^n \left[ D(x_j) - A - B \cos \omega \sum_{i=1}^m a_i P_i(x_j) \right]^2. \quad (3)$$

Элементы информационной матрицы Фишера выражаются как

$$I_{kl} = \left(\frac{\omega B}{\sigma}\right)^2 \sum_{j=1}^n P_k(x_j) P_l(x_j) \sin^2 \omega \sum_{i=1}^m a_i P_i(x_j) = \frac{1}{2} \left(\frac{\omega B}{\sigma}\right)^2 \sum_{j=1}^n P_k(x_j) \times \\ \times P_l(x_j) \left[ 1 - \cos 2\omega \sum_{i=1}^m a_i P_i(x_j) \right]. \quad (4)$$

Оценим значения элементов  $I_{kl}$  для линейного описания аргумента:

$$P_1(x) = 1, P_2(x) = x. \quad (5)$$

При более богатом представлении аргумента точность оценки результирующей фазы будет ниже. Заменяем суммирование интегрированием на интервале, симметричном относительно начала координат:

$$I_{kl} \cong \frac{1}{2\Delta} C \int_{-d}^d P_k(x) P_l(x) [1 - \cos 2\omega (a_1 + a_2 x)] dx. \quad (6)$$

Здесь  $\Delta = 2d/n$  — дискретность съема данных;  $C = \left(\frac{\omega B}{\sigma}\right)^2$ ;

$$I_{11} = C \frac{n}{2} \left[ 1 - \frac{1}{\omega a_2 n \Delta} \sin \omega a_2 n \Delta \cos 2\omega a_1 \right]; \quad (7)$$

$$I_{12} = I_{21} = C \sin 2\omega a_1 \left[ \frac{1}{(2\omega a_2)^2 \Delta} \sin \omega a_2 n \Delta - \frac{n}{4\omega a_2} \cos \omega a_2 n \Delta \right];$$

$$I_{22} = C \frac{n^3 \Delta^2}{24} \left\{ 1 - \cos 2\omega a_1 \left[ \frac{6}{(\omega a_2 n \Delta)^2} \cos \omega a_2 n \Delta + \left( \frac{3}{\omega a_2 n \Delta} - \frac{6}{(\omega a_2 n \Delta)^3} \right) \times \right. \right. \\ \left. \left. \times \sin \omega a_2 n \Delta \right] \right\}.$$

Уже при рассмотрении пяти периодов интерферограммы  $\omega a_2 n \Delta = 5 \cdot 2\pi$  можно считать

$$I_{11} \cong C \frac{n}{2}; \quad I_{12} = I_{21} \cong 0; \quad I_{22} \cong C \frac{n^3 \Delta^2}{24};$$

при этом погрешность не превысит 10%. С увеличением же числа обрабатываемых циклов строгость оценки возрастает. В этом случае предельная точность оценивания коэффициентов  $a_1$  и  $a_2$  определяется выражениями:

$$\sigma_{a_1}^2 = \frac{2\sigma^2}{\omega^2 B^2 n}; \quad \sigma_{a_2}^2 = \frac{24\sigma^2}{\omega^2 B^2 (n\Delta)^2 n}.$$

Очевидно, что точность оказывается пропорциональной отношению сигнал/шум  $B/\sigma$  и объему обрабатываемого массива  $n$ . Естественно также, что она может быть повышена переходом к более высокой частоте интерферометра  $\omega$ . Менее очевидным, хотя часто интуитивно ощущаемым является пропорциональность точности оценки коэффициента при линейном члене квадрату числа обрабатываемых периодов  $\omega n \Delta$ . Практическое использование этого положения заключается в том, что опорный и информационный потоки направляются на негатив под разными углами.

В порядке иллюстрации рассмотрим один из алгоритмов, реализующих такой подход, на несколько упрощенной математической модели. Будем считать, что работа осуществляется на линейном участке характеристической функции фотоматериала, аргумент периодической функ-

ции меняется параболически, рабочая точка и уровень сигнала и шума на всем интервале сканирования постоянны:

$$I(x) = A + B \sin(a_0 + a_1x + a_2x^2) + \xi(x).$$

Целью будет являться определение «полной» фазы или «набега» фазы, т. е. оценка параметров  $a_i$ . Параметры  $A$  и  $B$  в такой постановке можно отнести к сопутствующим или мешающим. Оценки  $A$ ,  $B$ ,  $a_0$ ,  $a_1$ ,  $a_2$  можно искать по методу наименьших квадратов. Связь искомых параметров с исходными данными нелинейна, и непосредственное аналитическое решение не просматривается, поэтому приходится, выбрав нулевые приближения для всех параметров, использовать затем итерационную процедуру для их уточнения. Освободиться от оценивания  $A$  и  $B$ , не представляющих для нас интереса, при таком алгоритме не удается. В качестве нулевого приближения  $A$  и  $B$  можно использовать статистики

$$\hat{A} = \frac{1}{n} \sum_{i=1}^n I(x_i); \quad \hat{B} = \frac{\pi}{2n} \sum_{i=1}^n |I(x_i) - \hat{A}|,$$

основанные на замене  $I(x)$  синусоидой.  $\frac{1}{T} \int_0^T |\sin t| dt = \frac{2}{\pi}$ , если  $T$  кратно  $\pi$ .

Для получения первого приближения параметров  $a_i$  можно использовать информацию, содержащуюся в переходах через нуль функции

$$I(x_i) - \hat{A}.$$

Обозначим эти точки  $\tau_1, \dots, \tau_k$ . При отсутствии шума они должны удовлетворять системе уравнений

$$\begin{vmatrix} 1 & \tau_1 & \tau_1^2 \\ 1 & \tau_2 & \tau_2^2 \\ \cdot & \cdot & \cdot \\ 1 & \tau_k & \tau_k^2 \end{vmatrix} \begin{vmatrix} a_0 \\ a_1 \\ a_2 \end{vmatrix} = \pi \begin{vmatrix} 1 \\ 2 \\ \cdot \\ k \end{vmatrix}$$

Погрешности измерений делают систему несовместной, и оценки параметров находятся по методу наименьших квадратов. Определение точек  $\tau_i$  осложняется наличием шума, приводящего к многократному пересечению нулевого уровня. Чтобы устранить эту неопределенность, фиксируем момент входа  $t_q$  и выхода  $t_l$  сигнала из зоны пересечений. В качестве оценки  $\tau_i$  берем их среднее

$$\tau_i = \frac{t_q + t_l}{2}$$

при условии, что  $I(t_q) \cdot I(t_l) < 0$ . В противном случае переход не регистрируется. Такой способ оценивания переходов функции через нулевой уровень позволяет работать при сравнительно высоком уровне шума. Дальнейшее повышение помехоустойчивости достигается посредством предварительного скользящего параболического сглаживания. Моделирование подтвердило работоспособность алгоритма при уровне шума  $\sigma/B = 1,5 \div 2$ . При этом параметры сигнала имели значения:  $A=0$ ;  $B=3$ ;  $a_0=1$ ;  $a_1=0,12$ ;  $a_2=0,12 \cdot 10^{-2}$ . В обработку вводилось 100 измерений ( $x=1, \dots, 100$ ). Повышение шума выше указанного уровня выводило алгоритм на локальный минимум, и возникала необходимость в принудительном переборе для поиска удовлетворительного первого приближения параметров. Сбой алгоритма в данном случае были обусловлены пропуском одной или нескольких пар переходов сигнала через нулевой уровень либо регистрацией таких переходов при их отсутствии. Если

процент таких сбоев невелик, можно дополнить алгоритм таким образом, чтобы он обнаруживал эти сбои и нейтрализовал их действие. В том или ином облике сбои будут присущи любому алгоритму обработки сигналов с циклической функцией преобразования. Вторым неприятным качеством такого рода сигналов является пороговость работы обслуживающих их алгоритмов. Порог может быть различен у алгоритмов различной сложности и достоинства, но присутствует он всегда. В процессе моделирования алгоритма выяснилось, что основная часть машинного времени поглощается итерационной процедурой. Вместе с тем обеспечиваемое ею уточнение оценок, по сравнению с первым приближением, оказывается не столь существенным. Видимо, в большинстве практических случаев можно будет ограничиться построением оценок параметров по принципам, положенным в основу поиска нулевого приближения. А именно: не прибегая к оцениванию мешающих параметров, обрабатывать информацию, заключенную в моментах перехода сигнала, нормированного по амплитуде, через ряд фиксированных уровней.

В этом случае получаем алгоритм, инвариантный в некоторых пределах и к виду характеристической функции фотоматериала. Его достоинством будет сравнительная простота.

До сих пор речь шла об алгоритмах обработки данных интерферограмм, полученных в результате сканирования их по некоторому направлению. Можно использовать двумерное сканирование. При этом и объем массива обрабатываемых данных и количество оцениваемых параметров возрастают квадратично. В этом резерв повышения точности по сравнению с обработкой одномерных массивов.

Интересный класс задач связан с определением линий разрыва сигнала или его производных [7]. Здесь к числу мешающих параметров можно отнести не только характеристики фоторегистраторов, но и параметры в областях, где сигнал  $Z(x, y)$  непрерывен. Часто целью интерферометрического исследования является определение только этих линий аномального поведения сигнала.

#### ЛИТЕРАТУРА

1. Т. Хуанг, В. Шрейбер, О. Третьяк. Обработка изображений.— ТИИЭР, 1971, т. 59, № 11.
2. М. П. Гришин, Ш. М. Курбашов, В. П. Маркелов, В. А. Соколов. Обработка интерферограмм на ЦВМ.— Тезисы конференции «Автоматизация научных исследований на основе применения ЭВМ». Новосибирск, 1972.
3. А. М. Фаянс, Л. П. Ярославский. Некоторые вопросы обработки и расшифровки интерферограмм на ЦВМ.— Тезисы конференции «Автоматизация научных исследований на основе применения ЭВМ». Новосибирск, 1972.
4. Б. М. Пушной. К теории измерительных систем с периодическими функциями преобразования.— Автометрия, 1972, № 4.
5. Б. Л. Ван дер Варден. Математическая статистика. М., Изд-во иностр. лит., 1960.
6. Ю. В. Динник. Метод наименьших квадратов и основы теории обработки наблюдений. М., Физматгиз, 1962.
7. В. С. Кирячук, Б. Н. Луценко. Выделение ступенчатых изменений сигнала и его производных.— Тезисы конференции «Автоматизация научных исследований на основе применения ЭВМ». Новосибирск, 1972.

*Поступила в редакцию 13 октября 1972 г.*