

УДК 535.411.854

В. И. Гужов

(Новосибирск)

**РАСШИРЕНИЕ ОБЛАСТИ ФАЗОВОЙ ОДНОЗНАЧНОСТИ
ПРИ ИНТЕРФЕРЕНЦИОННЫХ ИЗМЕРЕНИЯХ**

Предложен новый метод устранения фазовой неоднозначности, основанный на использовании серии измерений при различных ценах интерференционных полос. Метод позволяет резко увеличить диапазон интерференционных измерений без потери точности.

Одним из основных недостатков интерференционных измерительных систем является ограниченный диапазон измерений. Причина этого – периодический характер зависимости интенсивности от разности фаз интерферирующих волновых фронтов. Значения фаз могут быть однозначно определены только в пределах некоторого периода. Это ограничение, известное под названием фазовой или 2π -неоднозначности, присуще всем фазовым методам измерений.

В данной работе предложен новый метод, основанный на использовании серии измерений при различных ценах интерференционных полос, который позволяет резко увеличить диапазон измерений без потери точности [1–5]. Цена полосы определяется разностью оптического хода волн, при которой интерференционные полосы меняются на период. Ниже описание метода основано на использовании результатов нескольких интерференционных измерений при разных длинах волн.

Метод может быть использован для измерения разности фаз волновых фронтов, у которых значения в соседних точках отличаются более чем на длину волны. Отследить подобные изменения с помощью существующих алгоритмов определения полной фазы не удастся.

Для практического использования метода необходимо провести анализ погрешностей и разработать методику коррекции возникающих ошибок. Использование метода без коррекции ошибок невозможно, поскольку при переходе от непрерывного множества измеряемых величин к целочисленным значениям возникают ошибки, учет которых традиционными методами анализа погрешностей затруднителен. Поэтому цель данной статьи – анализ ошибок и разработка быстрых алгоритмов коррекции.

Изменение фазы на 2π для волн с разными периодами происходит при различных значениях оптической разности хода. С помощью интерференционных методов можно определить только оптические разности хода волн (ОРХ) в пределах длины волны источника освещения. Добавляя к

найденным результатам значения, кратные длинам волн, получим набор предполагаемых решений, соответствующих любой из используемых волн. Если найденные значения не совпадают, они отбрасываются, поскольку физически измеряемая разность хода не меняется. Только совпадающие для всех используемых волн значения могут претендовать на роль действительных ОРХ.

Таким образом, можно составить таблицу решений, но более удобно найти аналитическое решение.

Каждому целому числу отвечает определенный остаток от деления на целое положительное m , которое называется модулем. Если двум целым a и b отвечает один и тот же остаток r , то они называются равноостаточными по модулю m . Сравнимость записывается как

$$a \equiv b \pmod{m}. \quad (1)$$

Числа, сравнимые по модулю m , образуют класс чисел по модулю m . Рассмотрим систему сравнений с одним неизвестным, но с разными и притом попарно простыми модулями:

$$\begin{aligned} x &\equiv b_1 \pmod{m_1}, \\ x &\equiv b_2 \pmod{m_2}, \\ &\dots \\ x &\equiv b_k \pmod{m_k}. \end{aligned} \quad (2)$$

Решение системы сравнений запишется в виде

$$X_0 = M_1 N_1 b_1 + M_2 N_2 b_2 + \dots + M_k N_k b_k, \quad (3)$$

где M_i и N_i находятся из теоремы об остатках [6]. Диапазон измерений, при котором решение единственно, определяется произведением $m_1 m_2 \dots m_k$.

Для определения полной фазы при использовании двух длин волн необходимо найти решение целочисленной системы сравнений:

$$\text{ОРХ} \equiv \text{ОРХ}_1 \pmod{m_1}, \quad (4a)$$

$$\text{ОРХ} \equiv \text{ОРХ}_2 \pmod{m_2}, \quad (4b)$$

где ОРХ_1 и ОРХ_2 – остатки от деления искомой разности хода на соответствующую длину волны; m_1 и m_2 – соответствующие каждой из волн целые взаимно простые числа.

Максимальный диапазон, в пределах которого достигается единственность решения, определяется выражением

$$D = \lambda_1 m_2 = \lambda_2 m_1, \quad (5)$$

где λ_1 и λ_2 – используемые длины волн; m_1 и m_2 – соответствующие им модули.

В отличие от двухчастотного метода [7] выбор волн не произволен. Он определяется следующими факторами: точностью измерительной системы и взаимной простотой модулей. Величина модуля зависит от точности измерений исходных значений. Например, при измерительной точности, обеспе-

чивающей два десятичных разряда, источнику освещения с длиной волны $\lambda_1 = 0,633 \text{ мкм}$ можно поставить в соответствие модуль $m_1 = 63$, длине волны $\lambda_2 = 0,529 \text{ мкм}$ – $m_2 = 53$. Использование метода возможно, если m_1 и m_2 – взаимно простые числа (например, 53 и 63).

При экспериментальных измерениях фазового поля точность зависит от используемого аналого-цифрового преобразователя. При изменении от 0 до 2π квантованные значения фазы меняются от 0 до 255 (для наиболее часто используемого 8-разрядного АЦП). Поскольку необходимо перейти от фазовых к целочисленным значениям, изменяющимся в пределах $0 + m_i$, возникают ошибки, связанные с квантованием исходных величин. В этом случае прямое использование алгоритма невозможно: необходимо учитывать возникающие погрешности.

В силу специфики предложенного метода анализ ошибок имеет свои особенности. Рассмотрим систему сравнений с $m_1 = 53$ и $m_2 = 63$. Решение системы имеет следующий вид:

$$X \equiv 1008b_1 + 2332b_2 \pmod{3339}. \quad (6)$$

Из таблицы решений (рис. 1) видно, что при ошибке даже на единицу младшего разряда в b_1 или b_2 ошибка определения полной фазы может быть весьма значительной.

Соседние значения, расположенные по диагоналям, возрастают на единицу. Числа меняются от 0 до $\min(m_1, m_2, \dots, m_k)$ по главной диагонали, затем последовательно возрастают по другим диагоналям. При одиночной ошибке (при определении единственной координаты b_i) числа в таблице по этой координате отличаются на $M_i N_i$, взятые по модулю $m_1 m_2 \dots m_k$.

Если погрешность задания чисел по всем координатам одинакова и равна k , то вокруг истинного результата находится окрестность с радиусом k , в которой расположены числа с заведомо неправильными значениями. Назовем эту область окрестностью грубых промахов. На рис. 1 выделены диагонали, вдоль которых значения последовательно возрастают. Если ограничить диапазон изменения результата, то в указанную окрестность могут попадать только значения, не входящие в диапазон измерения. Видно, что при некотором заданном диапазоне изменения результата окрестность определяется величинами ΔX_1 и ΔX_2 , для нахождения которых достаточно получить их в первой строке и в первом столбце.

Если минимальные расстояния между соседними диагоналями ΔX_1 и ΔX_2 таковы, что границы окрестностей грубых промахов не пересекаются,

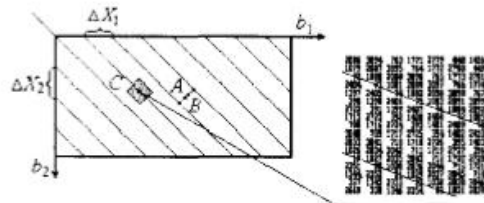


Рис. 1. Часть таблицы решений системы сравнений с модулями $m_1 = 53$ и $m_2 = 63$

то появляется возможность компенсации ошибок. Для этого находится ближайшее число, расположенное на одной из соседних допустимых диагоналей. Естественно, что ошибка, связанная с неопределенностью положения числа на достоверной диагонали, остается, но она не превышает погрешности определения исходных значений.

Для определения размеров окрестности грубых промахов сигнала необходимо найти все значения, попадающие в допустимый диапазон в нулевой строке и нулевом столбце, затем расположить их в порядке возрастания и вычислить минимальное расстояние между двумя ближайшими значениями. Поскольку числа по диагоналям отличаются на единицу, то допустимые значения расположены на них.

На рис. 1 отмечены диагонали, на которых находятся допустимые значения при ограничении диапазона измерения полной фазы величиной $D = 300$. При этом ограничении минимальное расстояние между диагоналями равно 10. Вокруг каждого числа, попадающего в заданный диапазон, образуется окрестность, состоящая из чисел, превышающих этот диапазон. Окрестность грубых промахов для точки с координатами $b_1 = 10, b_2 = 23$ на рис. 1 заштрихована. Видно, что если абсолютная ошибка при определении b_i не превышает ± 5 , то для рассматриваемого случая она может быть компенсирована.

Для коррекции ошибочных значений можно предложить следующий алгоритм. По координатам b_1 и b_2 находим решение системы сравнений. Если это значение не попадает в выбранный диапазон, вычисляем все решения последовательно в окрестности размером 1, 2, ... до тех пор, пока не получим удовлетворяющее нас решение. Понятно, что максимальная величина окрестности будет равна размеру окрестности грубых промахов. Этот способ требует множества вычислительных операций, среди которых наиболее медленной является операция умножения.

Нами разработано несколько быстрых алгоритмов. Все они основаны на сведении n -мерной задачи к $(n - 1)$ -мерной. Анализ окрестности для двумерного случая заменяется анализом значений, расположенных на некоторой прямой. Вместо поиска подходящих значений в трехмерном кубе, проводится анализ в одной из выбранных плоскостей. Кроме того, быстрые алгоритмы не используют операций умножения. Схема первого быстрого алгоритма показана на рис. 2.

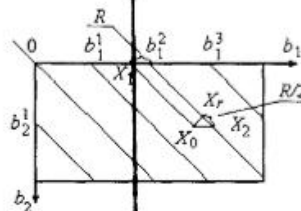


Рис. 2. Схема быстрого алгоритма, основанного на массивах допустимых значений

Пусть X_0 – точка, для которой решение системы сравнений не попадает в заданный диапазон. Для определения ближайшей к ней точки X_r , удовлетворяющей необходимым условиям, требуется:

1. Найти X_1 – точку пересечения с осью $0 - b_1$ или с осью $0 - b_2$. В этом случае координаты точки X_0 будут $(b_1 - b_2, 0)$ или $(0, b_2 - b_1)$.

2. Определить координаты ближайшего к X_0 допустимого значения b'_j для каждого из модулей, отсортированных в порядке возрастания. Нахождение координат заключается в выборке допустимых значений из массива координат.

3. Определить координаты допустимого значения X_r . Для этого находится точка X_2 , удовлетворяющая диапазону допустимых значений. Затем для нахождения более точного значения добавляется или вычитается $R/2$ в зависимости от того, где (слева или справа от X_0) расположена ближайшая допустимая диагональ, здесь R – расстояние между X_1 и координатой допустимой точки b'_j .

Однако этот метод эффективен только для двумерного случая. Для случая трех сравнений необходимо иметь $2m_3$ массивов допустимых значений. При больших значениях m_3 это приведет к существенным затратам памяти.

Более медленным, но простым и универсальным является второй предлагаемый нами алгоритм (рис. 3).

Пусть X_0 – точка с координатами b_1, b_2 , для которой решение системы сравнений не попадает в заданный диапазон. Выберем произвольную координату b_1 или b_2 и с последовательно увеличивающимся шагом $1, 2, \dots$ будем анализировать точки слева и справа от X_0 . Для нахождения решений достаточно учесть, что точки слева и справа будут отличаться на величину $M_x N_x$, взятую по модулю $m_1 m_2 \dots m_k$ (3), т. е. необходимо добавлять или вычитать это значение и определять значение по модулю. После определения значения, попадающего в заданный диапазон (например, X_2), производится уточнение. Для этого находится координата точки X_r , которая лежит внизу на расстоянии $R/2$, если найденное значение расположено справа, или вверх допустимой диагонали, если найденное значение лежит слева от точки X_0 .

В этом алгоритме также отсутствует операция умножения, но необходимы дополнительные операции сложения и взятия по модулю.

Для трехмерного случая поиск допустимого значения производится вдоль плоскости (рис. 4). Для n -мерного случая поиск осуществляется на $(n-1)$ -мерном кубе.

Коррекция ошибочных результатов приводит к уменьшению динамического диапазона измерений на величину, равную размеру окрестности

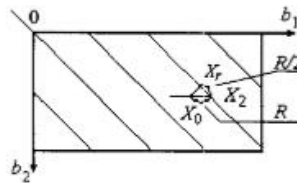


Рис. 3. Схема универсального алгоритма

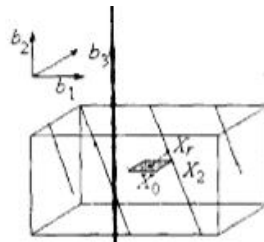


Рис. 4. Схема универсального алгоритма для трех модулей

грубых сбоев. Допустим, что значение модулей определяется некоторой величиной m . Для двумерного случая в таблице имеется $m \times m$ чисел. Если окрестность грубых сбоев равна m , то максимальный диапазон измерений становится равным m^2/n . Поскольку в каждой диагонали m чисел, то для двумерного случая число периодов, для которых значение полной фазы определяется однозначно, составляет m^2/n или m/n . Например, для $m_1 = 53$, $m_2 = 63$ при $n = 10$ число возможных периодов 5 (вместо возможных 53). Если увеличить точность измерений, то диапазон увеличивается пропорционально числу значимых знаков. Например, для $m_1 = 531$, $m_2 = 631$ и $n = 7$ диапазон составит около 76 периодов. Для трех модулей ($m_1 = 53$, $m_2 = 63$, $m_3 = 73$) число периодов равно 250, а при $m_1 = 531$, $m_2 = 631$, $m_3 = 731$ максимальное число периодов составит уже 2600 периодов с окрестностью грубых сбоев, равной 10. В этом случае возможный диапазон измерений охватывает от 0,005 мкм до 1 мм.

Дальнейшее расширение диапазона измерений возможно при повышении точности или при использовании еще одного добавочного измерения. Более эффективно увеличение точности [5]. В этом случае возможно восстановление полной фазы без коррекции. Диапазон измерений при этом максимален и пропорционален произведению модулей.

Компьютерное моделирование алгоритма осуществлялось с помощью задания эталонных фазовых фронтов (клин, сфера). Вначале определялся максимальный диапазон измерений, затем вычислялись остатки от деления

значений фронта в каждой точке плоскости на ряд взаимно простых модулей. Остатки масштабировались от 0 до 255. Полученные массивы чисел (аналоги реально измеряемых полей фаз) являлись исходными данными для восстановления по ним полной фазы. Эти массивы

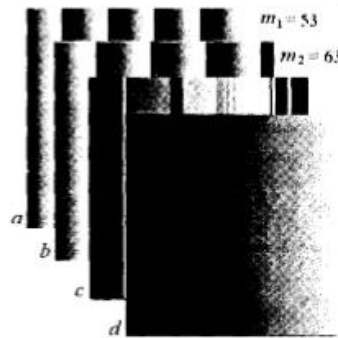
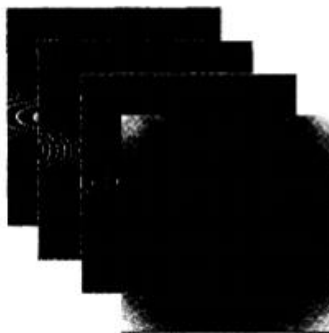


Рис. 5. Восстановление полной фазы по двум полям фаз:

a, b – исходные фазовые поля с $m_1 = 53$, $m_2 = 63$; c – поле фаз, восстановленное без коррекции ошибок; d – восстановленная целочисленным методом полная фаза с коррекцией

Рис. 6. Восстановление полной фазы сферического волнового фронта с диапазоном 200 периодов



соответствуют полям фаз, определенным по интерферограммам, с ценами полос, относящимися друг к другу как заданные модули.

Рассмотрим систему из двух сравнений с модулями $m_1 = 53$, $m_2 = 63$ (рис. 5). В этом случае решение определяется выражением (6). При диапазоне измерений $D = 300$ (пять периодов) диапазон окрестности грубых сбоев равняется 10. При таком диапазоне осуществляется точное восстановление полной фазы. При увеличении диапазона измерений размер окрестности сокращается до трех и не может быть уверенно скорректирован.

Как видно из рис. 5, с, восстановление фазы без коррекции невозможно, даже в случае идеальных фронтов. Это связано с ошибками квантования, рассмотренными в [5].

Для увеличения диапазона измерений необходимо использовать еще одно дополнительное измерение. Пусть $m_1 = 53$, $m_2 = 63$, $m_3 = 73$, тогда решение определяется выражением

$$X = 101178b_1 + 65773b_2 + 76797b_3 \pmod{243747}. \quad (7)$$

При заданных значениях модулей максимальный диапазон фазового фронта составит 250 периодов. В этом случае окрестность грубых сбоев равна 11.

Рассмотрим результаты моделирования сферического волнового фронта с диапазоном измерения полной фазы 10000 рад (200 периодов) (рис. 6, 7).

При использовании длины волны порядка 0,5 мкм 250 периодам соответствует 0,1 мм. Необходимо заметить, что исследуемый метод позволяет

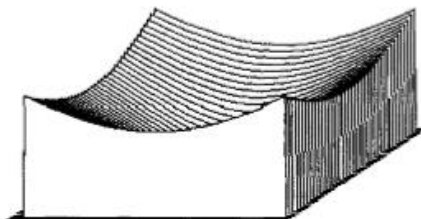


Рис. 7. Трехмерный график волнового сферического фронта с диапазоном 200 периодов

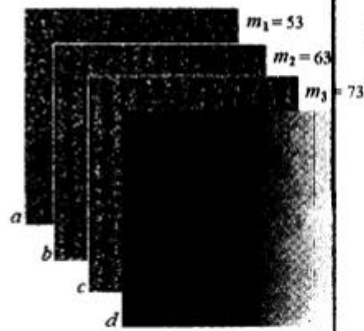


Рис. 8. Восстановление полной фазы по трем полям фаз с диапазоном изменения 20 периодов и перепадами в соседних точках более длины волны

восстанавливать полную фазу при перепадах в соседних точках, значительно превышающих длину волны. Рассмотрим, например, фазовый фронт, имеющий перепады в соседних точках более длины волны. На рис. 8–10 показан такой случай, когда при общем диапазоне изменения поля полных фаз, рав-

ном 1000, перепад в соседних точках составляет 100. Этот случай соответствует измерениям объектов с диффузной поверхностью.

Точность определения полей фаз методами пошагового сдвига может составлять сотую часть используемой длины волны [7]. Если диапазон измеренной фазы при вводе в компьютер составляет 8 разрядов, то цена

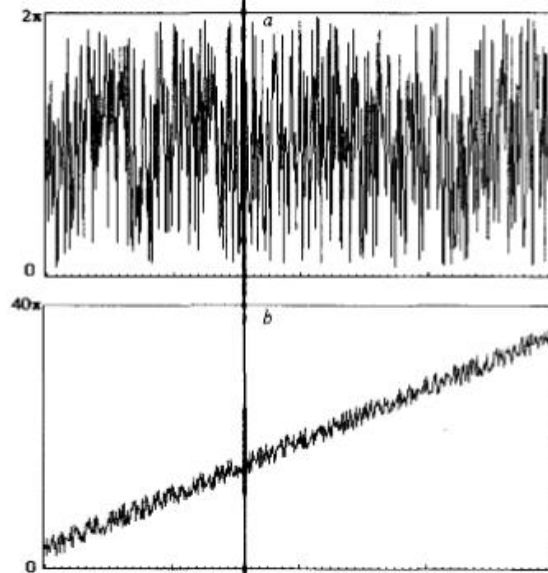


Рис. 9. Графики сечений по центральной строке фазового поля: *a* – сечение по центральной строке исходного фазового поля, определенного по модулю $m_1 = 53$; *b* – сечение поля восстановленной полной фазы

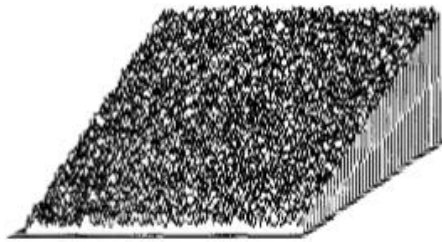


Рис. 10. Трехмерный график волнового фронта (см. рис. 8, д)

полосы (следовательно, и отношение между выбранными модулями чисел) может задаваться с точностью, составляющей не менее двух десятичных знаков. Этим объясняется выбор значений модулей при моделировании.

Метод не налагает дополнительных требований к виду фазового поля и может использоваться при анализе волновых фронтов, значительно отличающихся от опорных. В этом случае становится возможным применение интерференционных методик для анализа волновых фронтов, отраженных от объектов с диффузной поверхностью, что позволяет разрабатывать измерительные системы с характеристиками, необходимыми для измерения деталей в процессе производства без предварительной шлифовки.

СПИСОК ЛИТЕРАТУРЫ

1. Гужов В. И., Солодкин Ю. Н. Использование свойств целых чисел для расшифровки интерферограмм // Оптика и спектроскопия. 1988. 65, вып. 5. С. 1123.
2. Gushov V. I., Solodkin Yu. N. Automatic processing of fringe patterns in integer interferometers // Optics and Lasers Eng. 1991. 14. P. 311.
3. Гужов В. И., Солодкин Ю. Н. Оценка точности целочисленного интерферометра // Оптика и спектроскопия. 1988. 65, вып. 6. С. 1313.
4. Гужов В. И., Солодкин Ю. Н. Анализ точности определения полной разности фаз в целочисленных интерферометрах // Автометрия. 1992. № 6. С. 24.
5. Гужов В. И. Практические аспекты измерения фазы в интерферометрии // Автометрия. 1995. № 5. С. 25.
6. Виноградов И. М. Основы теории чисел. М.: Наука. 1972.
7. Creath K. Step height measurement using two-wavelength phase-shifting interferometry // Appl. Opt. 1987. 26, N 14. P. 2810.

Поступила в редакцию 11 октября 1997 г.