

В. И. Гужов, Е. В. Картавых
(Новосибирск)

**ПРОБЛЕМА ФАЗОВОЙ НЕОДНОЗНАЧНОСТИ
И ЕЕ РЕШЕНИЕ В ЛАЗЕРНОЙ ИНТЕРФЕРОМЕТРИИ**

Рассмотрен метод устранения фазовой неоднозначности, основанный на использовании серии измерений при различных ценах интерференционных полос и позволяющий резко увеличить диапазон интерференционных измерений без потери точности. Нахождение решения основано на целочисленной арифметике. Поскольку начальные фазовые значения задаются с некоторой погрешностью, необходима коррекция ошибочных значений. Приведен быстрый алгоритм коррекции для произвольного числа измерений. Возможности метода показаны на экспериментальном примере восстановления полной фазы при измерении угла между интерферирующими волнами.

Введение. Периодический характер зависимости интенсивности интерферограммы от фазовой разности интерферирующих волновых фронтов приводит к тому, что фазовые значения могут восстанавливаться однозначно только в пределах периода. Период полос определяется разностью фаз, при которой регистрируемая интенсивность меняется от минимума до максимума. Величина периода (цена интерференционной полосы) зависит от длины волны источника освещения, угла между интерферирующими волновыми фронтами и коэффициента пропускания среды.

Использование существующих методов устранения фазовой неоднозначности, основанных на пространственном развертывании поля фаз, позволяет восстановить фазовое распределение по полю в ограниченном диапазоне (порядка десятков полос). Расширение диапазона при увеличении цены полосы вызывает рост погрешности измерений.

Проблема фазовой неоднозначности возникает не только в оптической интерферометрии, но и в других фазовых измерениях (в радиодиапазоне или в акустике). Ее решение позволит значительно расширить область применения интерференционных методов и использовать интерференционные оптические системы для исследования объектов сложной формы с диффузной поверхностью без наличия дополнительных компенсирующих элементов.

В данной работе рассмотрено устранение фазовой неоднозначности в лазерной интерферометрии, однако выводы и рекомендации без нарушения общности могут быть перенесены на другие фазовые методы.

Описание метода. Целочисленный метод устранения фазовой неоднозначности, основанный на использовании серии измерений при различных ценах интерференционных полос, позволяет резко увеличить диапазон измерений без потери точности [1–4].

Изменение фазы на 2π при измерениях с разными ценностями полос происходит при различных значениях оптической разности хода интерферирующих волн. С помощью интерференционных методов можно определить только оптические разности хода волн в пределах цены полосы. Добавляя к полученным результатам значения, кратные ценам полос, получим набор возможных решений. Если найденные значения не совпадают, они отбрасываются, поскольку физически измеряемая разность хода не меняется.

Можно составить таблицу решений, просчитав все возможные сочетания, но более удобно найти аналитическое решение, воспользовавшись китайской теоремой об остатках, известной в теории чисел [5].

Каждому целому числу отвечает определенный остаток от деления на целое положительное m , называемое модулем. Сравнимость записывается как $x \equiv b_i \pmod{m_i}$. Числа, сравнимые по модулю m , образуют класс чисел по модулю m . Если провести несколько измерений с разными и притом попарно простыми модулями, получим систему сравнений с одним неизвестным x . Решение системы сравнений запишется в виде

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

где M_s и N_s находятся из условий $m_1 m_2 \dots m_k = M_s m_s$, $M_s N_s \equiv 1 \pmod{m_s}$. Диапазон, при котором решение единственно, определяется произведением модулей $m_1 m_2 \dots m_k$.

Максимальный диапазон, при котором удается устранить фазовую неоднозначность, определяется абсолютной величиной используемых модулей и числом исходных измерений. Чем точнее известны значения модулей, тем в большем диапазоне можно восстановить фазу.

Коррекция ошибочных значений. Рассмотрим в качестве примера систему двух сравнений с модулями $m_1 = 53$ и $m_2 = 63$. Решение системы имеет следующий вид:

$$X \equiv 1008b_1 + 2332b_2 \pmod{3339}.$$

Все возможные решения системы можно представить в виде таблицы. В этой таблице координаты по строкам и столбцам – значения остатков от 0 до m_1 и от 0 до m_2 соответственно.

Из анализа таблицы решений (рис. 1) видно, что даже при ошибке на единицу младшего разряда в b_1 или b_2 ошибка определения полной фазы может быть весьма значительной [4].

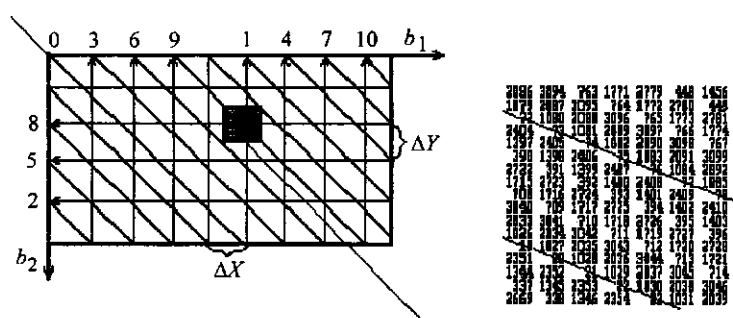


Рис. 1. Таблица решений системы с двумя сравнениями: $m_1 = 53$ и $m_2 = 63$

Соседние значения, расположенные вдоль диагоналей, возрастают на единицу. На рис. 1 показаны первые десять диагоналей, на которых расположены числа в порядке возрастания и выделенная часть этой таблицы.

Если погрешность задания данных 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. Из рисунка видно, что если ограничить диапазон изменения результата, то в окрестность могут попадать только значения, выходящие за пределы диапазона. При некотором заданном диапазоне изменения окрестность определяется величинами ΔX , ΔY . Эти значения зависят от выбранного ограничения диапазона и могут быть определены для конкретных значений модулей.

Если минимальные расстояния между соседними диагоналями таковы, что границы окрестности грубых промахов не пересекаются, появляется возможность компенсации ошибок. Для этого находится ближайшее число, расположенное на одной из ближайших допустимых диагоналей.

Необходимость коррекции ошибочных результатов приводит к уменьшению максимально возможного диапазона. Для восстановления полной фазы размер окрестности грубых промахов должен соответствовать погрешности при измерении исходных данных (чем больше погрешность, тем больше размер окрестности).

Например, для $m_1 = 53, m_2 = 63$ и окрестности с размером $r = 10$ возможна ошибка при определении исходных значений ± 5 . Это соответствует допустимой относительной ошибке при измерении оптической разности хода в пределах длины волны $\sim 7\%$. В этом случае диапазон составит только 5 периодов вместо возможных 53.

При $m_1 = 531, m_2 = 631$ и $r = 30$ допустимая ошибка в определении исходных данных $\sim 2,3\%$, диапазон ~ 19 периодов. При тех же модулях и $r = 7$ возможная ошибка $\sim 0,5\%$, диапазон ~ 80 периодов.

Для трех модулей ($m_1 = 53, m_2 = 63, m_3 = 73$) при размере окрестности 10 допустимая ошибка $\sim 6,8\%$, диапазон 250 периодов.

При $m_1 = 531, m_2 = 631, m_3 = 731$ допустимая ошибка $\sim 4\%$, диапазон 2600 периодов. При цене полосы 0,5 мкм максимальный диапазон оптического хода волн, который может быть измерен в этом случае, составит более 1,3 мм. При этом погрешность измерений не должна превышать $\lambda/25$.

Определение ближайшего допустимого значения перебором всех решений в окрестности по формуле (3) требует большого объема вычислительных ресурсов. При дискретизации интерферограммы на 512×512 точек и числе модулей 3 такая операция занимает несколько часов на процессоре "Pentium-II" с тактовой частотой 300 МГц.

Для практического использования метода необходимо разработать быстрые алгоритмы коррекции ошибочных значений. В работе [4] приведены алгоритмы для двух- и трехмерных случаев. Ниже рассматривается быстрый алгоритм для произвольного числа модулей.

Быстрый алгоритм коррекции ошибочных значений. Если соединить продолжения диагоналей при последовательном возрастании чисел в таблице решений (см. рис. 1), можно заметить, что при склейке верхней и нижней горизонтальных строк и левого и правого столбца образуется тор (рис. 2). Числа из таблицы решений расположены на пересечении меридианов и параллелей, показанных на рисунке. Видно, что числа последовательно возрастают по спирали, нанесенной на поверхность тора.

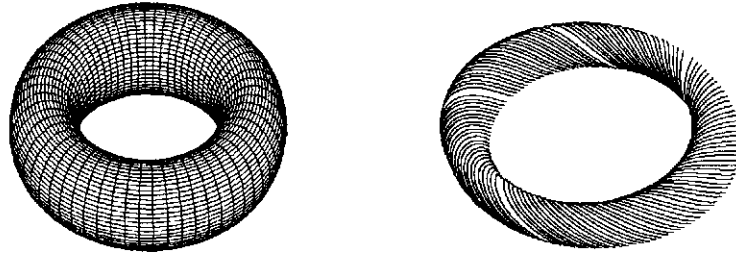


Рис. 2. Тор, образующийся в результате склейки таблицы решений: справа – последовательное возрастание чисел на первых десяти диагоналях

В данном случае задача определения числа по его модулярному представлению сводится к одномерной. Это вытекает из следующих соображений.

Разрежем тор вдоль одного из меридианов. Тогда он превращается в круговой цилиндр с двумя краевыми окружностями. Закрепим неподвижно одну окружность и станем закручивать цилиндр так, чтобы вторая окружность сделала k оборотов. Всякая прямолинейная образующая цилиндра при этом обратится в винтовую линию, обходящую ось цилиндра k раз. Если снова склеить оба края, то получим топологическое отображение тора на самого себя. При таком отображении параллели тора превращаются в винтообразные кривые и наоборот [6].

Таким образом, если хранить одномерный массив значений на окружности, перпендикулярной образующей тора, можно восстановить все значения k -й таблицы раскручиванием по винтовой линии. Для определения всех значений в этом сечении достаточно хранить один массив AX размером m_1 , в котором записаны решения системы при $b_k = 0, \dots, b_3 = 0, b_2 = 0$.

Этот алгоритм обобщается на многомерный случай. Его структура имеет следующий вид:

- 1) $v_{1,1} = b_1 - b_k \bmod(m_1); v_{2,1} = b_2 - b_k \bmod(m_2); \dots; v_{k-1,1} = b_{k-1} - b_k \bmod(m_{k-1});$
- 2) $v_{1,2} = v_{1,1} - v_{k-1,1} \bmod(m_1); v_{2,2} = v_{2,1} - v_{k-1,1} \bmod(m_2); \dots$
 $\dots; v_{k-2,2} = v_{k-2,1} - v_{k-1,1} \bmod(m_{k-2});$
 $s = s + (M_1 N_1 + M_2 N_2 + \dots + M_{k-1} N_{k-1}) v_{k-1,1} \bmod(m_1 m_2 \dots m_k);$
- 3) $v_{1,3} = v_{1,2} - v_{k-2,2} \bmod(m_1); v_{2,3} = v_{2,2} - v_{k-2,2} \bmod(m_2); \dots$
 $\dots; v_{k-3,3} = v_{k-3,2} - v_{k-2,2} \bmod(m_{k-3});$
 $s = s + (M_1 N_1 + M_2 N_2 + \dots + M_{k-2} N_{k-2}) v_{k-2,2} \bmod(m_1 m_2 \dots m_k);$
 \dots
- $k-1$) $v_{1,k-1} = v_{1,k-2} - v_{2,k-2} \bmod(m_1);$
 $s = s + (M_1 N_1 + M_2 N_2) v_{2,k-2} \bmod(m_1 m_2 \dots m_k);$
- k) $X = AX[v_{1,k-1}] + s + b_k \bmod(m_1 m_2 \dots m_k).$

В данном алгоритме отсутствуют операции умножения и очень просто выполняется операция взятия модуля, которая может быть приведена к последовательности вычитаний.

Использование предложенного алгоритма позволяет на несколько порядков увеличить быстродействие при коррекции ошибочных результатов.

Экспериментальное моделирование метода. Для экспериментальной проверки алгоритма использовался интерферометр Тваймана – Грина (рис. 3).

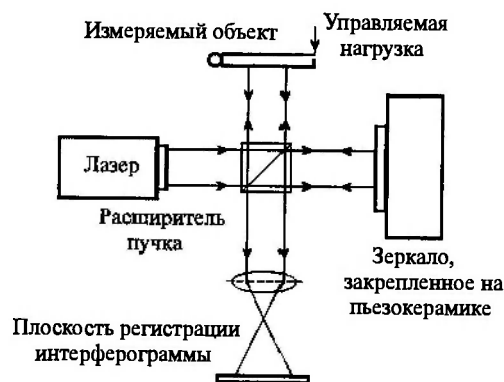


Рис. 3. Интерферометр Тваймана – Грина с изменяемым углом между интерферирующими пучками

В опорном плече интерферометра размещается зеркало, закрепленное на пьезокерамике, для задания управляемого фазового сдвига между интерферирующими пучками. Плоское зеркало в другом плече изменяет угол между интерферирующими пучками. Один из концов зеркала закреплен на шарнире, другой нагружен управляемой от компьютера пьезокерамикой. Механическая конструкция закрепления зеркала обеспечивает хорошую повторяемость изменения угла.

На рис. 4 показаны серии интерферограмм с фазовыми сдвигами $0, \pi/2, \pi, 3\pi/2$. Проводились две серии измерений при изменении угла между интерферирующими пучками. Поля фаз по модулю 2π определялись методом пошагового фазового сдвига. Погрешность при определении фазовых значений составляла $\lambda/50$.

Для восстановления полной фазы использовался целочисленный алгоритм. Поскольку точное значение изменения угла определить сложно, находился средний размер полос для каждой серии измерений и ему ставилось в соответствие значение модуля, используемого для расшифровки. Для приведенных серий интерферограмм взаимно простые значения модулей составляют $m_1 = 547$ и $m_2 = 630$. Решение системы сравнений для этих значений модулей примет вид: $X = 91350b_1 + 253261b_2 \text{ mod}(344610)$.

При ограничении интервала изменений полной фазы до 15 периодов диапазон окрестности грубых сбоев равен 15. Это значит, что допускается ошибка в определении исходных фазовых значений ± 7 .

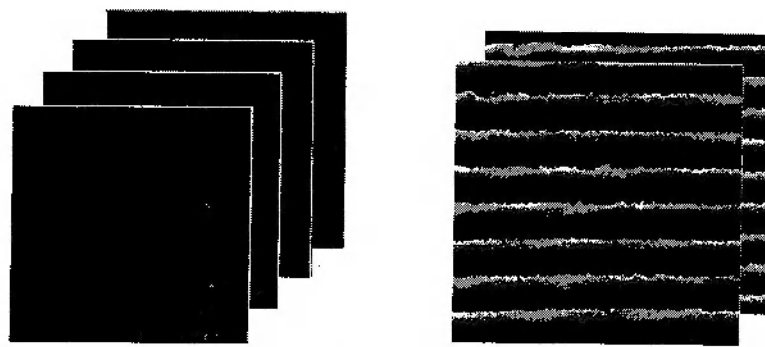
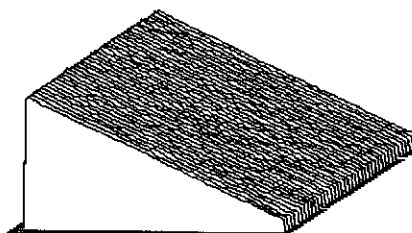


Рис. 4. Интерферограммы и поля фаз

Рис. 5. Восстановленное поле полной фазы



Восстановленное поле полной фазы показано на рис. 5.

Точное изменение угла сдвига на заданную величину достаточно сложно обеспечить. Поэтому более предпочтительно использовать источник освещения с переменной длиной волны, поскольку длина волны когерентного излучения может задаваться с достаточной точностью.

Заключение. Целочисленный метод устранения фазовой неоднозначности значительно расширяет область фазовой однозначности. При этом увеличение динамического диапазона не связано с уменьшением точности.

Погрешность измерений фазовых значений в пределах периода приводит к необходимости коррекции ошибочных значений. Коррекция становится возможной при ограничении допустимого диапазона.

Метод не налагает дополнительных ограничений на вид фазового поля и может использоваться при анализе волновых фронтов сложной формы. Становится возможным применение интерференционных методик для анализа объектов с диффузной поверхностью, что позволяет разрабатывать измерительные системы с характеристиками, необходимыми для практического 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 in 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. Расширение области фазовой однозначности при интерференционных измерениях // Автометрия. 1998. № 3. С. 99.
5. Кнут Д. Искусство программирования для ЭВМ. М.: Мир, 1977. Т. 2.
6. Гильберт Д., Кон-Фоссен С. Наглядная геометрия. М.: Наука, 1981.

Новосибирский государственный
технический университет,
E-mail: vig@ref.nstu.nsk.su

Поступила в редакцию
20 апреля 2000 г.