

УДК 531.715

В. И. Гужов

*(Новосибирск)***ПРАКТИЧЕСКИЕ АСПЕКТЫ ИЗМЕРЕНИЯ ФАЗЫ
В ИНТЕРФЕРОМЕТРИИ**

Рассматриваются два локальных метода удаления фазовой неоднозначности: метод двухдлинноволновой интерферометрии с управляемым фазовым сдвигом и целочисленный метод, основанный на модульной арифметике.

Введение. Все интерференционные методы измерений обладают фазовой неоднозначностью, являющейся следствием периодичности изменения интерференционной картины. Это приводит к тому, что значения фаз могут определяться только в пределах периода.

Наиболее распространенным способом устранения 2π -неоднозначности являются методы развертывания фаз, основанные на анализе пространственной структуры фазового поля. Если измеряемое поле фаз плавное, перепады в соседних точках не превышают π , и можно проследить пространственную границу переходов через период, то нанесением рассекающих векторных линий между соседними зонами с замкнутыми полосами и использованием априорной информации о порядке полос проблема фазовой неоднозначности может быть решена. При анализе пространственного распределения фаз в некоторых областях вблизи границ переходов к значениям в предыдущей зоне добавляется или вычитается 2π . Такие методы накладывают существенные ограничения на вид измеряемых фазовых фронтов. Кроме того, они очень чувствительны к накоплению ошибок, так как информация о предыдущих восстановленных точках используется для определения значений в следующих точках.

Вследствие этого большой интерес вызывают локальные методы, позволяющие по результатам наблюдения в одной точке расширять диапазон, в котором фазовая неоднозначность отсутствует, без анализа пространственной структуры поля. В предлагаемой статье рассмотрены два метода: метод двухдлинноволновой интерферометрии и целочисленный метод.

Основой для устранения фазовой неоднозначности являются фазовые значения в пределах периода, полученные из интерферограмм при различных значениях периодов интерференционных полос.

Определение фазовых значений в пределах периода. Для оценки 2π -неоднозначности исходными являются фазовые значения, полученные способом внесения управляемого фазового сдвига (phase-shifting interferometry — PSI), для каждой из двух длин волн λ_1 и λ_2 . При использовании нескольких волн с различными длинами наиболее удобно использовать вычисления по методу, предложенному Carré [1]. В отличие от других алгоритмов не требуется предварительной калибровки фазовых сдвигов при изменении длины волны. Требуется только, чтобы значения сдвигов были постоянными.

Регистрируется четыре интерферограммы, каждая из которых получена с определенным сдвигом фаз между объектной и опорной волной. Если фазовый

сдвиг увеличивать на 2α при регистрации каждой из последующих интерферограмм, получим

$$A(x, y) = I_0(x, y) [1 + \gamma(x, y) \cos(\varphi(x, y) - 3\alpha)],$$

$$B(x, y) = I_0(x, y) [1 + \gamma(x, y) \cos(\varphi(x, y) - \alpha)],$$

для интерферограмм с фазовыми сдвигами -3α , $-\alpha$, α , 3α соответственно; $\varphi(x, y)$ — значения разности фаз между опорным и объектным пучками; $I_0(x, y)$ — средняя интенсивность; $\gamma(x, y)$ — модуляция интенсивности.

Фазовые значения в каждой точке (x, y) вычисляются по формуле

$$\varphi = \operatorname{arctg} \frac{\sqrt{[(A - D) + (B - C)][3(B - C) - (A - D)]}}{\{(B + C) - (A + D)\}}. \quad (2)$$

Функция arctg определена от $-\pi/2$ до $\pi/2$. Диапазон неоднозначности может быть расширен до $(-\pi, \pi)$ анализом знаков выражений

$$(B - C) = (2I_0\gamma \sin\alpha) \sin\varphi, \quad (3)$$

$$(B + C) - (A + D) = (2I_0\gamma \cos\alpha \cdot \sin^2\alpha) \cos\varphi. \quad (4)$$

Метод двухдлинноволновой интерферометрии с управляемым фазовым сдвигом. Наиболее известным из двух рассматриваемых методов является метод двухдлинноволновой интерферометрии с управляемым фазовым сдвигом (TWPSI — two-wavelength phase-shifting interferometry) [2—5].

При вычислении разности фаз, измеренных при двух волнах с разными длинами, получается фазовый фронт с периодом, равным эквивалентной длине волн λ_{eq} . Оптическую разность хода волн (optical path difference — OPD) в точке (x, y) можно определить из следующих выражений:

$$\varphi_1 + 2\pi n_1 = (2\pi/\lambda_1) \text{OPD}, \quad (5)$$

$$\varphi_2 + 2\pi n_2 = (2\pi/\lambda_2) \text{OPD},$$

где φ_1 и φ_2 — фазы, измеренные для волн с длинами λ_1 и λ_2 ; n_1 и n_2 — целые числа длин волн, соответствующие одному значению оптической разности хода. Вычитая одно равенство из другого, получим

$$\varphi_1(x, y) - \varphi_2(x, y) + 2\pi n_{\text{eq}} = (2\pi/\lambda_{\text{eq}}) \text{OPD}(x, y), \quad (6)$$

где

$$\lambda_{\text{eq}} = \lambda_1 \lambda_2 / |\lambda_1 - \lambda_2|. \quad (7)$$

Таким образом, выбором величины λ_{eq} определяется диапазон измерения, в котором отсутствует фазовая неоднозначность. При использовании двух волн 488 и 514 нм λ_{eq} равна 9,47 мкм, при $\lambda_1 = 633$ и $\lambda_2 = 529$ нм λ_{eq} равна 3,22 мкм. Использование существующих мод в аргоновом и гелий-неоновом (He—Ne) лазерах позволяет изменять эквивалентную длину волны от 2 до 50 мкм.

Целочисленный метод. Другой метод расширения однозначности при фазовых измерениях основан на целочисленной арифметике [6—9]. Изменение фазы на 2π для каждой из волн происходит при различных значениях оптиче-

ской разности хода между опорной и объектными волнами. Поэтому в общем случае одной разности хода соответствуют разные фазовые значения при использовании волн с различными частотами. При пересчете по полученным фазам можно определить OPD в пределах периода. Добавляя к полученным результатам значения, кратные периодам, получим набор решений, соответствующих каждой длине волны. Если соответствующие результаты из наборов не совпадают, они отбрасываются. Только совпадающие значения могут претендовать на роль действительных OPD [10].

Графическая интерпретация этого метода представлена на рис. 1. Исходными данными являются результаты измерений OPD при двух волнах с длинами λ_1 и λ_2 . Поскольку результаты фазовых измерений определяются в каждой точке с точностью до периода соответствующей волны, для каждой из волн существует набор решений, который состоит из вертикально расположенных точек, отстоящих друг от друга на длину волны. Совокупность значений, удовлетворяющих всем наборам, определяет область нахождения искомого OPD. Из рисунка видно, что совокупность значений, удовлетворяющих обоим наборам, повторяется с некоторым периодом. Этот период определяет диапазон однозначного нахождения фазы при заданных значениях длин волн.

Можно составить таблицу решений, просчитав все возможные сочетания, но более удобно найти аналитическое решение. Это можно сделать, воспользовавшись теоремой об остатках, известной в теории целых чисел [11].

Ищется решение целочисленной системы сравнений

$$\begin{aligned} \text{OPD} &\equiv \text{OPD}_1(\text{mod } m_1), \\ \text{OPD} &\equiv \text{OPD}_2(\text{mod } m_2), \end{aligned} \quad (8)$$

где знак (\equiv) обозначает сравнение; OPD_1 и OPD_2 — остатки от деления искомой разности хода на соответствующую длину волны; m_1 и m_2 — соответствующие каждой из волн целые взаимно простые числа. Они определяются из соотношения

$$m_2\lambda_1 = m_1\lambda_2. \quad (9)$$

В данном случае волне с длиной λ_1 ставится в соответствие целое число 5, волне с длиной λ_2 — число 3. По теореме об остатках область, в которой решение единственно, определяется произведением этих чисел и решение имеет вид

$$\text{OPD} \equiv 6\text{OPD}_1 + 10\text{OPD}_2(\text{mod } 15). \quad (10)$$

Динамический диапазон, в пределах которого достигается единственность решения, определяется числом целых длин волн

$$\lambda_{\text{int}} = \lambda_1 m_2 = \lambda_2 m_1. \quad (11)$$

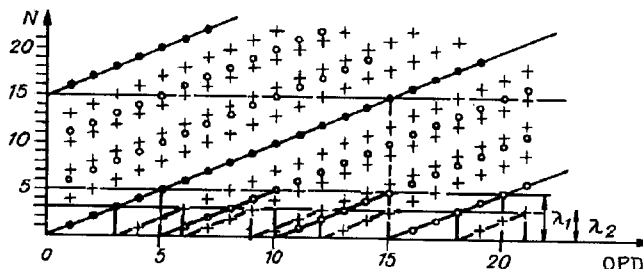


Рис. 1. Измеренные в пределах периода оптические разности хода и набор значений, удовлетворяющих измеренным значениям

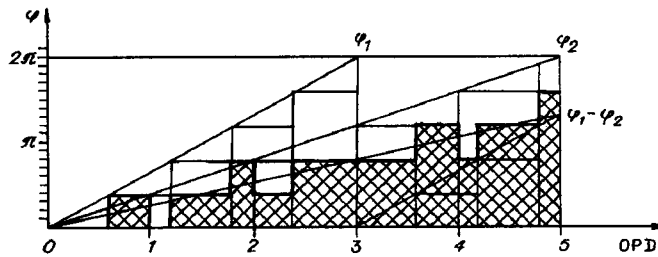


Рис. 2. Разность фаз двух пилообразных функций после квантования с одинаковыми уровнями фазы

Пусть имеется две волны $\lambda_1 = 0,633$ и $\lambda_2 = 0,529$ мкм. Если соответствующим периодам могут быть поставлены в соответствие целые взаимно простые числа 633 и 529, то диапазон, в котором отсутствует фазовая неоднозначность, равен 334,857 мкм, т. е. 633 полосы с периодом λ_2 .

Влияние погрешностей при измерении фаз на диапазон определения полной фазы. В TWPSI анализируются разности фаз. Однако в отличие от классического метода двухдлинноволновой голографической интерферометрии вычитаются квантованные значения фаз. Поэтому существенное влияние начинают оказывать ошибки, связанные с квантованием фазовых значений.

Рассмотрим случай, когда фазовые значения определяются с одной и той же точностью для разных волн (рис. 2). В этом случае максимальная ошибка при вычислении разности стремится к $2\pi/N$, где N — число дискрет при измерении фазовых значений. При этом, как видно из рисунка, нарушается монотонность функции разности.

Эффект несогласованного квантования OPD сглаживается, если величина кванта определяется в зависимости от соответствующего периода значения OPD (рис. 3). В этом случае максимальная ошибка определения разности составляет:

$$\Delta(\varphi_1 - \varphi_2) = (2\pi\lambda_2)/(N\lambda_{\text{eq}}) = (2\pi/N)(\lambda_2/\lambda_{\text{eq}}). \quad (12)$$

При согласованном с размером OPD квантовании эта ошибка меньше, сохраняется монотонность функции разности. Однако в сложившейся практике измерений обычно используется квантование фазы на одинаковые уровни фазы, поэтому необходимо учитывать ошибки от подобного квантования OPD при разных длинах волн.

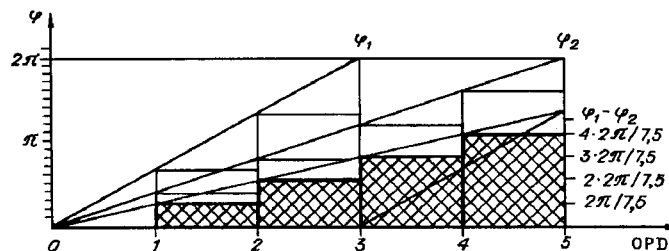


Рис. 3. Разность фаз двух пилообразных функций после согласованного квантования

Рассмотрим рис. 1. В этом случае $m_1 = 5$, $m_2 = 3$ и фазовые значения при согласованном с ОРД квантовании разбиваются на $N_1 = 5n$ и $N_2 = 3n$ уровней. При $n = 1$ имеем

$$\varphi_1 = 0, 2\pi/3, 2\pi \cdot 2/3, 0(2\pi), 2\pi/4, 2\pi \cdot 2/4, \dots, \quad (13)$$

$$\varphi_2 = 0, 2\pi/5, 2\pi \cdot 2/5, 2\pi \cdot 3/5, 2\pi \cdot 4/5, 0(2\pi), 2\pi/5, \dots \quad (14)$$

Разность фаз $\varphi_1 - \varphi_2$, взятая по модулю 2π (рис. 4), равна

$$\varphi_1 - \varphi_2 = 0, 2\pi/7,5, 2\pi \cdot 2/7,5, \dots, 2\pi \cdot 8/7,5, \dots, 2\pi \cdot 14/7,5, 0(2\pi), \dots \quad (15)$$

Из рис. 4 видно, что разность фаз с периодом λ_{eq} нарастает от 0 до 2π при изменении ОРД от 0 до $\lambda_{\text{eq}} = 7,5$, при этом число точек, приходящихся на период, равно $N_{\text{eq}}n \pm 1$, где N_{eq} — целая часть от $m_1 m_2 / [m_1 - m_2]$. Можно заметить, что начало второго периода в (15) $2\pi \cdot 8/7,5$ точно не совпадает с $2\pi n$. Абсолютно точные значения разности будут повторяться только через период, определяемый целочисленным методом и равный произведению $m_1 m_2$. На рис. 4 — это начало третьего периода. Если отличия больше чем $2\pi/N$, можно различать эти значения и таким образом расширять диапазон измерения ОРД.

В методе, основанном на модульной арифметике, анализируются значения оптической разности хода волн. Достигается большее, чем в TWPSI, расширение диапазона, в котором фазовые значения определяются однозначно, однако требования к точностным характеристикам при измерении фазовых значений выше.

Необходима абсолютно точная информация о соотношении двух длин волн. В действительности же при определении такого соотношения имеется некоторая погрешность. Это приводит к быстрому накоплению ошибок. Пусть $\Delta\lambda$ — погрешность от неправильного определения длины волны λ . Тогда из выражения

$$\Delta k > \lambda/Q, \quad (16)$$

где Q — число элементов квантования, можно определить k периодов волны, на которых будет накоплена ошибка на один элемент разрешения. При диапазоне измерения порядка 5 см с использованием He—Ne-лазера n равно 10^5 . В этом случае отношение $\Delta\lambda/\lambda$ должно быть $0,5 \cdot 10^{-7}$ при 256 уровнях квантования. При использовании методов стабилизации частоты можно получить относительную стабильность частоты и, следовательно, степень монохроматичности порядка 10^{-12} — 10^{-13} [12]. Наиболее распространенные лазеры имеют длину когерентности 30—50 см и относительную стабильность частоты

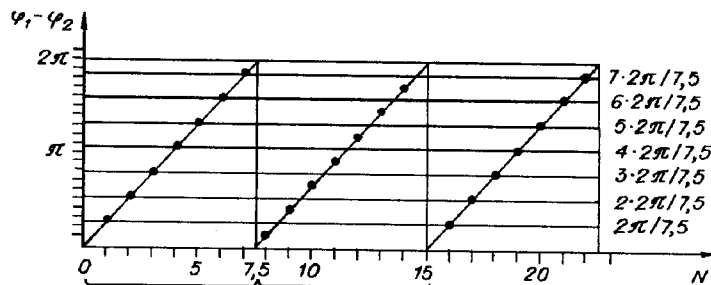


Рис. 4. Разность фаз двух пилообразных функций с периодами 5 и 3

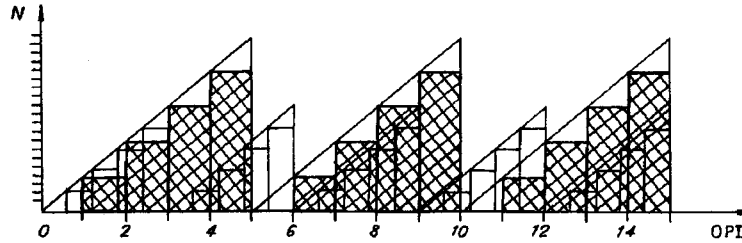


Рис. 5. Квантованные значения функций

10^{-6} — 10^{-7} . Этого достаточно для измерения фазового фронта в пределах нескольких сантиметров.

Целочисленный метод очень чувствителен к ошибкам, вызванным квантованием сигнала. Из рис. 5 видно, что погрешность, вызванная несогласованием с величиной OPD квантованием сигнала, приводит к ошибке на единицу младшего разряда. Если при использовании метода TWPSI ошибка при определении младшего разряда приводит к пропорциональному увеличению погрешности, то целочисленный метод дает неправильное значение определяемого периода. Такая ошибка может быть очень большой. Поэтому требования к точности определения фазовых значений в пределах периода выше. В [8] рассмотрено влияние одиночных ошибок и показано, что для коррекции ошибок на несколько единиц младшего разряда необходимо определить область устойчивости. Коррекция неточных результатов приводит к некоторому уменьшению динамического диапазона. Если точно известно соотношение длин волн, а точность определения фазы меньше, то динамический диапазон уменьшается пропорционально потере точности. Например, при значениях m_1 , m_2 , равных 633 и 529, пусть возможная точность определения фазы — 256 двоичных разрядов. Погрешность определения фазы составит ± 2 двоичных разряда, и при возможном расширении диапазона полной фазы до 633 полос с периодом λ_2 диапазон будет в 4 раза меньше.

Абсолютная погрешность определения OPD методом TWPSI равна

$$\Delta \text{OPD} = \Delta(\varphi_1 - \varphi_2)\lambda_{\text{eq}} / (2\pi) = 2\Delta\varphi\lambda_{\text{eq}} / (2\pi) = \Delta\varphi\lambda_{\text{eq}} / \pi. \quad (17)$$

Видно, что неопределенность при измерениях возрастает с ростом эквивалентной длины волны. Поэтому для повышения точности уточняют результаты, полученные методом TWPSI, значениями, полученными при отдельных измерениях с длинами волн λ_1 и λ_2 .

Диапазон двухдлинноволнового метода обратно пропорционален абсолютной разности используемых длин волн. Чем меньше эта разность, тем больше диапазон. Использование еще одного измерения с отличной от применяемых длиной волны не позволяет увеличить диапазон, поскольку новую волну необходимо выбрать близкой к λ_{eq} , что в большинстве случаев практически не реализуемо. Для целочисленного метода подобного ограничения не существует. Диапазон может быть значительно расширен использованием серии измерений при нескольких различных длинах волн.

Общими для обоих методов являются следующие замечания.

Измеряемое поле фаз может приводить к интерференционным картинам, в которых полосы имеют пространственные частоты, не разрешаемые детектором. В этом случае для увеличения пространственного разрешения применяют маску, состоящую из небольших отверстий, существенно меньших размера элемента детектора.

Необходимо учитывать эффекты, связанные с ахроматизмом оптических элементов. Если интерференционная картина меняет размер или пространственно сдвигается при изменении длины волны от хроматических aberrаций,

результатирующие фазы будут иметь ошибки, вызванные неправильным выбором соответствующих точек.

Заключение. Расширение диапазона измерений, в котором отсутствует фазовая неоднозначность при использовании целочисленного метода и метода TWPSI, определяется используемыми длинами волн.

Для расширения диапазона при использовании метода TWPSI необходимо выбирать источники освещения с близкими длинами волн. Однако неопределенность при измерениях возрастает с ростом эквивалентной длины волны.

При использовании целочисленного метода выбираются волны с длинами, отношение между которыми должно быть точно известно. При этом диапазон целочисленного метода может значительно превышать диапазон двухдлинно-волнового метода. В отличие от TWPSI можно резко увеличить диапазон использованием нескольких длин волн. Однако требования к точностным характеристикам измерительной системы повышаются.

СПИСОК ЛИТЕРАТУРЫ

1. Carre P. Installation et utilisation du comparateur photoelectrique et interferential du Bureau international des poids et mesures // Metrologia. 1966. 2(1).
2. Cheng Y.-Y., Wyant J. C. Two-wavelength phase shifting interferometry // Appl. Opt. 1984. 23. P. 4539.
3. Creath K., Cheng Y.-Y., Wyant J. C. Contouring aspheric surfaces using two-wavelength phase-shifting interferometry // Optica Acta. 1985. 32, N 12. P. 1455.
4. Creath K. Step height measurement using two-wavelength phase-shifting interferometry // Appl. Opt. 1987. 26, N 14. P. 2810.
5. Creath K., Wyant J. C. Direct phase measurement of aspheric surface contours // Proc. SPIE. 1986. 645. P. 101.
6. Гужов В. И., Солодкин Ю. Н. Использование свойств целых чисел для расшифровки интерферограмм // Оптика и спектроскопия. 1988. 65, вып. 5.
7. Gushov V. I., Solodkin Yu. N. Automatic processing of fringe patterns in integer interferometers // Opt. and Lasers Eng. 1991. 14. P. 311.
8. Гужов В. И., Солодкин Ю. Н. Оценка точности целочисленного интерферометра // Оптика и спектроскопия. 1988. 65, вып. 6.
9. Гужов В. И., Солодкин Ю. Н. Анализ точности определения полной разности фаз в целочисленных интерферометрах // Автометрия. 1992. № 6.
10. Greivenkamp J. E. Sub-Nyquist interferometry // Appl. Opt. 1987. 26, N 24. P. 5245.
11. Виноградов И. М. Основы теории чисел. М.: Наука, 1972.
12. Звелто О. Физика лазеров. М.: Мир, 1979.

Поступила в редакцию 9 февраля 1995 г.