

УДК 681.787: 681.518.5

Динамическое оценивание фазы в интерферометрии управляемого фазового сдвига методом адаптивной фильтрации Винера

© 2020 г. **И. П. Гуров, доктор техн. наук; И. А. Артемьева; В. О. Капанова**

Университет ИТМО, Санкт-Петербург

E-mail: gurov@mail.ifmo.ru

Поступила в редакцию 30.12.2019

DOI:10.17586/1023-5086-2020-87-05-42-53

Представлены результаты применения метода адаптивной фильтрации Винера для динамического оценивания параметров интерферометрических сигналов. Показана возможность непосредственного вычисления параметров интерференционных полос, в том числе фазы полос, через полученные коэффициенты адаптивного фильтра Винера по критерию минимума дисперсии ошибки. Рассматриваемый метод не вносит ограничений на количество и размер задаваемых фазовых сдвигов и позволяет определять фактическое значение фазового сдвига. Представлены оценки погрешностей метода в условиях воздействия шума наблюдений и случайных отклонений шага фазового сдвига.

Ключевые слова: оптический контроль, фаза интерференционных полос, динамическая обработка интерференционных полос, фильтрация Винера.

Коды OCIS: 070.2025, 120.3180, 120.5050

ВВЕДЕНИЕ

Методы интерферометрии управляемого фазового сдвига (ИФС) широко используются при решении задач высокоточного оптического контроля (см., например [1]). Предложены многочисленные алгоритмы обработки набора зарегистрированных интерференционных картин с заданным взаимным фазовым сдвигом для точного определения фазы интерференционных полос. Ввиду высокой чувствительности интерферометрические системы подвержены влиянию неточности значений фазового сдвига, воздействию шума и других внешних факторов [2, 3], вследствие чего предложены различные варианты алгоритмов обработки набора интерференционных картин с компенсацией погрешностей и повышенной помехоустойчивостью [1–6]. В так называемых

обобщенных методах ИФС используются произвольные значения фазового сдвига с вычислением фактического фазового сдвига для точного определения фазы [7, 8].

Алгоритмы, позволяющие определять значения фазы интерференционных полос в динамическом режиме, представляют особый интерес ввиду возможности оценивания объектов, подверженных малым отклонениям их формы, например, вследствие дрейфа температуры окружающей среды и влияния других внешних факторов в процессе измерений. В этой связи алгоритмы, в которых не вводится ограничение на размер и количество фазовых сдвигов, наиболее перспективны.

Динамическая обработка интерферометрических сигналов в ИФС может рассматриваться с позиций теории фильтрации в частотной

области, как это впервые показано в работе [9]. Фильтрация осуществляется в форме линейного суммирования с весовыми коэффициентами отсчетов интерферометрического сигнала в каждом пикселе по последовательности регистрируемых интерференционных картин. Набор коэффициентов определяет импульсный отклик фильтра и его частотную характеристику, вычисляемую с использованием преобразования Фурье (см., например [10]).

В традиционных алгоритмах ИФС значения и количество весовых коэффициентов фиксированы и не зависят от входных данных. Однако при таком подходе проявляется хорошо известная проблема, заключающаяся в наличии исчерпывающей априорной информации (или ее отсутствии) об оцениваемом полезном сигнале и шуме. Широко распространенный подход к оптимизации частотной характеристики фильтра в ИФС основывается на простейших априорных предположениях и состоит в обеспечении подавления фоновой составляющей сигнала на нулевой частоте и кратных гармоник по отношению к частоте интерференционных полос. Оптимизация обычно осуществляется соответствующим размещением нулевых значений модуля частотной характеристики фильтра [3, 9, 10]. Важно отметить, что при такой оптимизации фильтра в ИФС размер фазового сдвига, определяющий частоту интерференционных полос по последовательности интерференционных картин, должен быть заранее известен, и его возможные отклонения (детерминированные и случайные) должны учитываться при задании частотной характеристики фильтра. В частности, при использовании узкополосного фильтра, настроенного на предполагаемую частоту интерференционных полос (см., например [10]), проблема состоит в нелинейности фазочастотной характеристики фильтра, что вносит значительные погрешности фазы даже при весьма малых неконтролируемых отклонениях фактического значения частоты от предполагаемого значения.

Сравнительные оценки точности многих известных алгоритмов определения фазы интерференционных полос в ИФС, представленные в работе [11], позволили выделить наиболее предпочтительные варианты алгоритмов для частных случаев влияния линейных и квадратичных отклонений фазового сдвига,

нелинейности фотоприемника и кратных гармоник. Различия точностных характеристик алгоритмов отражают степень априорного знания (предположений), принятого в расчет при создании алгоритма, и степень соответствия предположений и реальных данных, получаемых в конкретном эксперименте.

Если априорная информация не полна и затруднительно сделать выбор наиболее подходящего алгоритма, целесообразно использовать адаптивные алгоритмы, автоматически настраивающиеся на наблюдаемые реальные данные согласно заданному критерию качества.

Рассматриваемые проблемы в задачах фильтрации были решены Н. Винером [12] при представлении полезных сигналов и шума как случайных процессов. Фильтр Винера (ФВ) основан на оптимизации спектральной плотности мощности на выходе фильтра применительно не к единственной реализации сигнала, но к ансамблю реализаций, относящихся к конкретному виду эксперимента. Известно, что спектральная плотность мощности стационарного случайного процесса связана с его корреляционной функцией через преобразование Фурье (теорема Винера–Хинчина). При этом частотную характеристику фильтра можно определить, принимая в расчет спектральную плотность мощности сигнала и шума, либо их корреляционные функции.

Данный подход был развит Р. Калманом, впервые распространившим фильтрацию случайных процессов на динамические системы с обработкой сигналов в пространстве состояний (см. оригинальную работу Р. Калмана [13]). Обработка интерферометрических сигналов в ИФС с помощью фильтров калмановского типа [14, 15] основывается на концепции рассмотрения параметров интерференционных полос как случайных величин. Такой подход позволяет построить алгоритмы с произвольными количеством и размером фазовых сдвигов, с обеспечением высоких помехоустойчивости и точности определения фазы, однако для этого требуется априорная информация о корреляционных матрицах шума и предполагаемых случайных отклонениях параметров интерференционных полос как начальных условий [15].

Важное преимущество рассматриваемых методов стохастической фильтрации состоит в их адаптивности, когда фильтр автоматиче-

ски настраивается на фактические характеристики сигнала в процессе конкретного эксперимента. В некоторых случаях, когда достаточно использовать ограниченную априорную информацию, возможно применение упрощенных алгоритмов, не требующих задания начальных условий. Одним из алгоритмов этого вида является адаптивный дискретный ФВ, общее описание которого приведено, например, в публикации [16], и его применение для решения задачи динамического оценивания фазы в ИФС, насколько нам известно, рассматривается в настоящей работе впервые.

Применение ФВ для определения фазы интерференционных полос в ИФС основывается на использовании не фиксированных, а динамически настраиваемых весовых коэффициентов фильтра согласно критерию минимума дисперсии ошибки по отношению к заданному опорному сигналу, содержащему априорную информацию о модели интерферометрического процесса. Как показано ниже, рассматриваемый метод не вносит ограничений на размер и количество фазовых сдвигов, обеспечивает возможность определения фактического значения фазового сдвига и высокую точность получаемых динамических оценок фазы.

ФИЛЬТР ВИНЕРА И ЕГО ПРИМЕНЕНИЕ В ИНТЕРФЕРОМЕТРИИ УПРАВЛЯЕМОГО ФАЗОВОГО СДВИГА

Широко используемая модель для набора интерференционных картин в ИФС определяется выражением

$$I(x, y, k) = I_a(x, y) \cos[\gamma k + \varphi(x, y)] + I_b(x, y), \quad (1)$$

где $I_a(x, y)$ — амплитудные значения интенсивности интерференционных полос, $I_b(x, y)$ — неинформативная фоновая составляющая интенсивности, $\varphi(x, y)$ — искомые значения фазы интерференционных полос, k — текущий номер интерференционной картины, $k = 0, \dots, K - 1$, γ — задаваемый фазовый сдвиг. Значения интенсивности (1) с помощью матрицы светочувствительных элементов преобразуются в фотоэлектрические сигналы $s = \mu I$, где μ — коэффициент преобразования, при этом модель (1) принимает вид

$$s(i, j, k) = a(i, j) \cos[\gamma k + \varphi(i, j)] + b(i, j), \quad (2)$$

где (i, j) обозначают пиксел с дискретными координатами $x_i = i\Delta x$, $y_j = j\Delta y$, Δx и Δy — шаги пикселей.

Рассмотрим скалярный интерферометрический сигнал в пикселе (i, j) по последовательности зарегистрированных интерференционных картин, который в соответствии с (2) определяется выражением

$$s(k) = a \cos(\gamma k + \varphi) + b, \quad (3)$$

где $\gamma = 2\pi/P$, P — количество шагов фазового сдвига на одном периоде интерференционных полос. В результате обработки сигнала (3) требуется получить точное значение фазы φ при априорно не известных значениях параметров a и b и отличии зарегистрированных значений сигнала от идеальной модели (3) вследствие влияния шума и других внешних факторов. Для получения статистических оценок сигнала, содержащего полезную составляющую и составляющую шума, возможно использовать ФВ.

Отсчеты сигнала $s(k)$ в ФВ порядка M взвешиваются с коэффициентами w_{m+1} , $m = 0, \dots, M - 1$ и суммируются. При этом сигнал $y(k)$ на выходе фильтра выражается в форме

$$y(k) = \sum_{m=0}^{M-1} w_{m+1} s(k-m). \quad (4)$$

Остаточную ошибку по отношению к заданному опорному сигналу $d(k)$ обозначим $\varepsilon(k) = y(k) - d(k)$. Функционирование адаптивного ФВ основывается на подстройке весовых коэффициентов w_{m+1} согласно критерию минимизации дисперсии отклонения от заданного сигнала

$$\langle \varepsilon^2(k) \rangle \rightarrow \min, \quad (5)$$

где угловые скобки обозначают усреднение по ансамблю реализаций сигнала.

Дисперсия ошибки (5) при векторном представлении отсчетов сигнала выражается в форме (см., например [16, 17])

$$\varepsilon^2(k) = d^2(k) - \mathbf{w}^T \mathbf{s}(k) \mathbf{s}^T(k) \mathbf{w} - 2d(k) \mathbf{s}^T(k) \mathbf{w}, \quad (6)$$

где $\mathbf{w} = [w_1 \ w_2 \ \dots \ w_M]^T$ — вектор весовых коэффициентов фильтра, $\mathbf{s}(k) = [s(k) \ s(k-1) \ \dots \ s(k-M-1)]^T$ — вектор значений сигнала. Вектор $\hat{\mathbf{w}}$, минимизирующий дисперсию ошиб-

ки (6), определяется условием равенства нулю градиента квадрата ошибки $\partial \varepsilon^2 / \partial \mathbf{w}$, которое выражается известным уравнением Винера–Хопфа в дискретной форме [16]

$$\mathbf{R}\hat{\mathbf{w}} = \mathbf{g}, \quad (7)$$

где $\mathbf{R} = \langle \mathbf{s}(k) \mathbf{s}^T(k) \rangle$ — корреляционная (не сингулярная) матрица компонентов сигнала $\mathbf{s}(k)$, имеющая обратную матрицу \mathbf{R}^{-1} , $\mathbf{g} = \langle d(k) \mathbf{s}(k) \rangle$ — вектор взаимной корреляции компонентов сигнала $\mathbf{s}(k)$ и опорного сигнала $d(k)$. Решение уравнения (7) выражается очевидным образом как

$$\hat{\mathbf{w}} = \mathbf{R}^{-1} \mathbf{g}. \quad (8)$$

Важно отметить, что в отличие от синтеза фильтра с детерминированной частотной характеристикой в традиционных алгоритмах ИФС ФВ представляет собой оптимальный фильтр согласно критерию (5), и определение весовых коэффициентов основывается на корреляционных свойствах сигнала, что обеспечивает применимость ФВ к ансамблю сигналов, принадлежащих случайному процессу. При этом ФВ учитывает случайные отклонения сигнала, влияние шума и искажений, которые определяют дисперсию ошибки (5) в процессе измерений при проведении конкретного эксперимента.

Обратимся к модели сигнала (3) и рассмотрим возможность применения ФВ для получения оценок параметров сигнала a , b и φ .

В методах ИФС количество регистрируемых интерференционных картин ограничено в целях сокращения времени измерений для минимизации влияния внешних факторов (помех). Данное требование соответствует условию минимума количества регистрируемых отсчетов K интерферометрического сигнала (3). Поскольку сигнал определяется тремя неизвестными параметрами a , b и φ , для их оценивания необходимо зарегистрировать не менее трех интерференционных картин (отсчетов сигнала (3) в каждом пикселе). Три отсчета сигнала ($M = 3$ в выражении (4)) складываются с весовыми коэффициентами w_1 , w_2 , w_3 и формируют выходной сигнал

$$y(k) = \sum_{m=0}^2 w_{m+1} s(k-m), \quad (9)$$

как это иллюстрируется схемой, приведенной на рис. 1. Полученные значения сигнала $y(k)$ сравниваются с заданным опорным сигналом $d(k)$, и весовые коэффициенты w_1 , w_2 , w_3 динамически подстраиваются с минимизацией отличия сигнала (9) от опорного сигнала.

Покажем, что ФВ позволяет не только минимизировать дисперсию ошибки (6), но и определить значения параметров входного сигнала, которые выражаются через значения коэффициентов фильтра.

Последовательность отсчетов сигнала (3) в схеме рис. 1 определяет выходной сигнал ФВ с учетом (9) в форме

$$\begin{aligned} y(k) = & w_1 [a \cos(\gamma k - \varphi) + b] + \\ & + w_2 [a \cos(\gamma(k-1) - \varphi) + b] + \\ & + w_3 [a \cos(\gamma(k-2) - \varphi) + b]. \end{aligned} \quad (10)$$

Выражение (10) преобразуется к виду

$$y(k) = p \cos(\gamma k + \varphi) + q \sin(\gamma k + \varphi) + c, \quad (11)$$

где коэффициенты p , q и c определяются выражениями

$$p = w_1 + w_2 \cos \gamma + w_3 \cos 2\gamma, \quad (12)$$

$$q = w_2 \sin \gamma + w_3 \sin 2\gamma, \quad (13)$$

$$c = (w_1 + w_2 + w_3)b. \quad (14)$$

Выражение (11) можно представить в форме

$$y(k) = r \cos(\gamma k + \varphi - \theta) + c, \quad (15)$$

где

$$r = \sqrt{p^2 + q^2}, \quad \cos \theta = p/r, \quad \sin \theta = q/r. \quad (16)$$

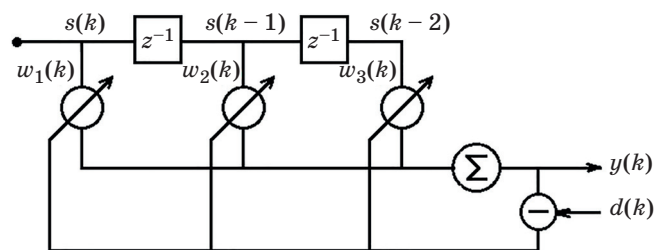


Рис. 1. Схема адаптивного фильтра Винера с тремя весовыми коэффициентами. Символ z^{-1} означает сдвиг на один шаг.

Комбинируя выражения (12), (13) и (16), получаем оценку амплитуды сигнала (15) в виде

$$r = a\sqrt{w_1^2 + w_2^2 + w_3^2 + 2(w_1w_2 + w_2w_3)\cos\gamma + 2w_1w_3\cos 2\gamma}. \quad (17)$$

Согласно (16)

$$\operatorname{tg}\theta = \frac{q}{p} = \frac{w_2 \sin\gamma + w_3 \sin 2\gamma}{w_1 + w_2 \cos\gamma + w_3 \cos 2\gamma}. \quad (18)$$

Априорно заданный опорный сигнал $d(k)$ (см. рис. 1) с известными амплитудой A и фоновой составляющей B определяется выражением

$$d(k) = A \cos\left(\frac{2\pi}{P}k\right) + B = A \cos(\gamma k) + B. \quad (19)$$

В соответствии с (15) минимизация отклонений от опорного сигнала (19) достигается при равенстве этих сигналов, а именно при условии

$$r \cos(\gamma k + \varphi - \theta) + c = A \cos(\gamma k) + B, \quad (20)$$

из которого следует, что $r = A$, $\varphi = \theta$, $c = B$. При этом получаем оценки амплитуды входного сигнала (3) (учитывая (17))

$$a = \frac{A}{\sqrt{w_1^2 + w_2^2 + w_3^2 + 2(w_1w_2 + w_2w_3)\cos\gamma + 2w_1w_3\cos 2\gamma}}, \quad (21)$$

фазы (в соответствии с (18))

$$\varphi = \operatorname{arctg} \frac{w_2 \sin\gamma + w_3 \sin 2\gamma}{w_1 + w_2 \cos\gamma + w_3 \cos 2\gamma} \quad (22)$$

и фоновой составляющей (согласно (14))

$$b = \frac{B}{w_1 + w_2 + w_3}. \quad (23)$$

Таким образом, с помощью ФВ возможно получение оценок амплитуды, фазы и фоновой составляющей интерферометрического сигнала (3) по формулам (21)–(23) с использованием коэффициентов $\hat{\mathbf{w}} = [\hat{w}_1 \hat{w}_2 \hat{w}_3]^T$ адаптивного ФВ. При этом точность определения параметров сигнала зависит от точности решения уравнения Винера–Хопфа (8). В этом уравнении корреляционная матрица \mathbf{R} и вектор взаимной корреляции \mathbf{g} представляют собой усредненные характеристики, вычисление которых не требует априорного знания о размере фазового сдвига γ в явном виде. Корреляционная матрица \mathbf{R} вычисляется только с использованием отсчетов регистрируемого интерферометрического сигнала.

Априорная информация автоматически учитывается в значениях вектора взаимной корреляции с заданным опорным сигналом (19).

Важно отметить принципиальное отличие ФВ от известного метода определения фазы гармонического сигнала при прямом вычислении взаимной корреляции входного и опорного сигналов [18].

Взаимная корреляционная функция сигналов (3) и (19) имеет вид (для упрощения полагаем в (19) $A = 1$, $B = 0$)

$$\begin{aligned} g(n) &= \langle d(k)s(k+n) \rangle = \\ &= \sum_{k=0}^{P-1} \cos(\gamma k) [a \cos[\gamma(k+n) + \varphi] + b] = \\ &= \sum_{k=0}^{P-1} [(a/2)\cos(2\gamma k + \gamma n + \varphi) + \\ &+ (a/2)\cos(\gamma n + \varphi) + b \cos(\gamma k)]. \end{aligned} \quad (24)$$

Второе слагаемое в выражении (24) содержит информацию о фазе φ (точнее, о взаимном сдвиге n , при котором $\varphi = -\gamma n$ и достигается максимум корреляции). Однако для получения этого значения необходимо устранить

первое и третье слагаемые за счет их усреднения по k точно в пределах периода, который априорно не известен. Другая проблема обусловлена требуемой малой погрешностью определения фазы: значения γn должны покрывать полный фазовый цикл 2π с высоким разрешением, чтобы найти оценку n точно в положении корреляционного максимума. Следовательно, необходимо зарегистрировать большое количество интерференционных картин P с малым взаимным фазовым сдвигом γ (по крайней мере порядка 100 картин, см. [18]), что на практике неприемлемо для методов ИФС, либо использовать дополнительные методы интерполяции.

При реализации рассматриваемого метода требуется регистрировать весьма малое количество интерференционных картин с произвольным фазовым сдвигом γ . Для повышения точности оценок элементов корреляционной матрицы \mathbf{R} (оценок их математического ожидания) целесообразно выполнить усреднение элементов матрицы по нескольким дополнительным отсчетам сигнала в форме

$$R(u, v) = \frac{1}{Q-2} \sum_{k=3}^Q s(k-u)s(k-v), \quad u, v = 1, 2, 3. \quad (25)$$

Выражение (25) соответствует усреднению произведений в $Q-2$ группах по три отсчета. Проведенные исследования показали, что высокая точность ФВ достигается при $Q=5$, т.е. для вычислений достаточно использовать пять последовательных отсчетов сигнала. Аналогичным образом вычисляется усредненный коэффициент взаимной корреляции входного и опорного сигналов в форме

$$g(v) = \frac{1}{Q-2} \sum_{k=3}^Q d(k)s(k-v), \quad v = 1, 2, 3 \quad (26)$$

при $Q=5$. Заметим, что при практической реализации алгоритма нормирующие множители $1/(Q-2)$ в выражениях (25) и (26) можно опустить, поскольку они входят и в левую, и в правую части уравнения (7).

Можно показать [16, 17], что достигаемая минимальная дисперсия ошибки ФВ (6) определяется выражением

$$\epsilon_{\min}^2 = d^2(k) - \mathbf{g}^T \hat{\mathbf{w}}. \quad (27)$$

Это выражение означает, что ФВ осуществляет компенсацию квадрата опорного сигнала суперпозицией взвешенных значений вектора взаимной корреляции входного и опорного сигналов, вычисленных при нескольких фазовых сдвигах (в нашем случае при $M=3$).

ПОГРЕШНОСТИ ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ ИНТЕРФЕРЕНЦИОННЫХ ПОЛОС ПРИ ИСПОЛЬЗОВАНИИ АДАПТИВНОГО ФИЛЬТРА ВИНЕРА

Исследования точности рассматриваемого метода проводились при оценивании среднего квадратичного отклонения (СКО) получаемых значений амплитуды, фоновой составляющей и фазы интерференционных полос в условиях воздействия аддитивного шума. Степень влияния шума устанавливалась отношением сигнал/шум (SNR), определяемым в форме

$$\text{SNR} = a / \sigma_n, \quad (28)$$

где a , как и ранее, — амплитуда сигнала, σ_n — СКО значений аддитивного гауссова шума с нулевым средним.

При моделировании фаза сигнала в формуле (2) задавалась согласно выражению

$$\varphi(i, j) = \eta[(i-i_0)^2 + (j-j_0)^2], \quad (29)$$

где (i_0, j_0) — дискретные координаты пиксела, в котором значение фазы равно нулю, $\eta = 2\pi[(L-i_0)^2 + (H-j_0)^2]^{-1}$ — нормирующий множитель, $L \times H$ — размеры интерференционной картины (в пикселах). Амплитуду и фоновую составляющую сигнала (2) принимали равными $a=1$, $b=1,5$ соответственно.

На рис. 2а иллюстрируется пространственное распределение фазы (29). На рис. 2б–г приведены значения весовых коэффициентов $\hat{\mathbf{w}}(i, j)$, полученные при пяти шагах фазового сдвига $\gamma = 2\pi/4$.

Погрешности определения параметров интерференционной картины оценивались в зависимости от SNR (28). Полученные оценки погрешностей представлены кривыми на рис. 3. Кривые получены при усреднении по 10^4 пикселах, соответствующим распределению фазы, показанному на рис. 2а.

Из рис. 3 видно, что метод обеспечивает получение малых погрешностей определения

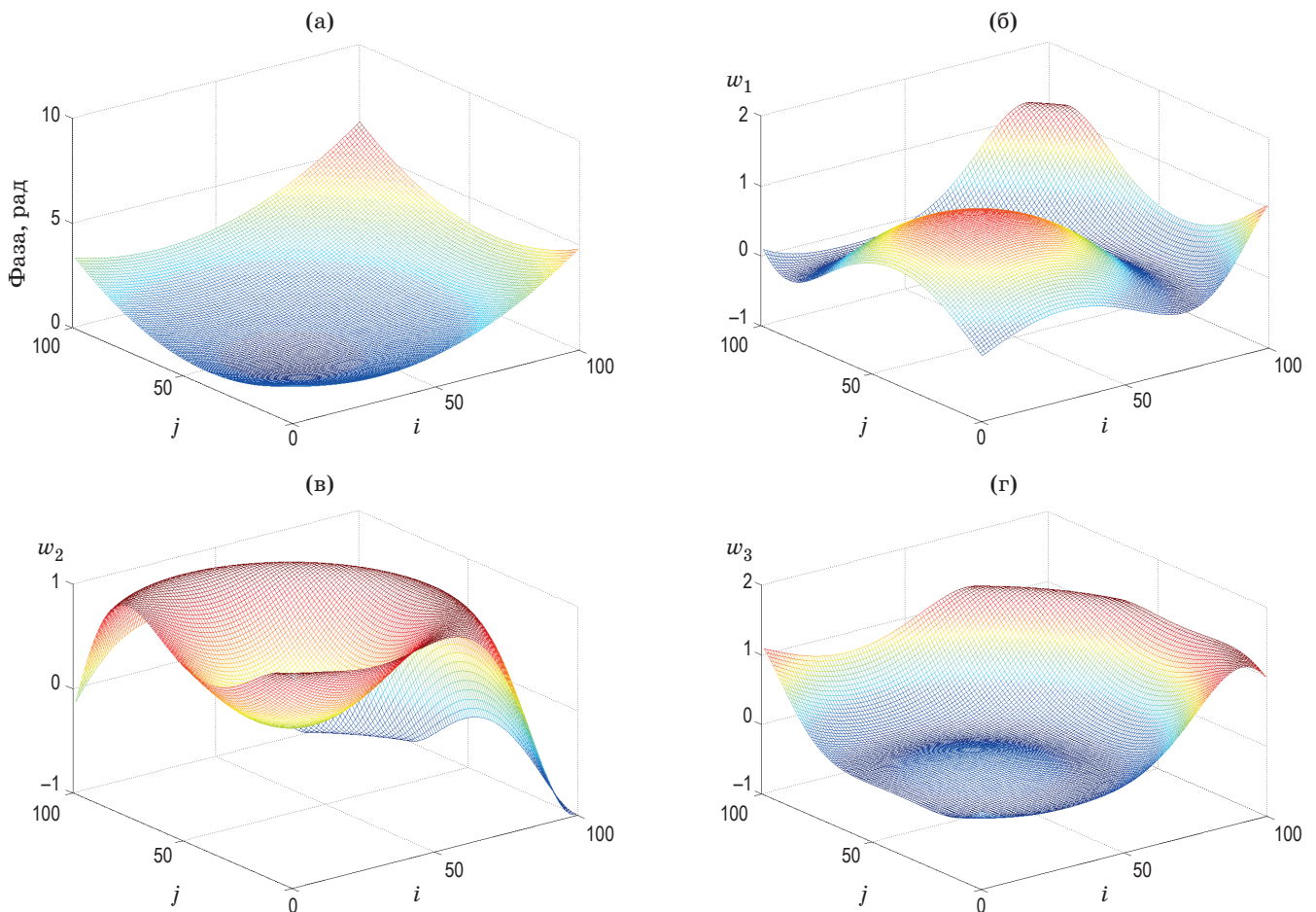


Рис. 2. Распределение фазы (29) форматом 100×100 пикселей при $i_0 = 30$, $j_0 = 40$ (а) и соответствующие значения весовых коэффициентов w_1 , w_2 , w_3 адаптивного фильтра Винера (б–г).

всех параметров интерференционной картины при отношении сигнал/шум выше 10. Отметим важное свойство: погрешность определения фазы интерференционных полос практически не зависит от задаваемого фазового сдвига, как это видно из кривых на рис. 3в.

Для получения приведенных выше оценок параметров интерференционных полос в рассматриваемом алгоритме достаточно использовать пять интерференционных картин ($Q = 5$ в формулах (25) и (26)) при произвольном фазовом сдвиге. Поскольку общее количество шагов фазового сдвига K в данном алгоритме не лимитируется, исследовалась эволюция получаемых оценок при «скользящей» обработке вновь поступающих отсчетов сигнала. Полученные результаты показали стабильность получаемых оценок параметров интерферометрических сигналов и снижение погрешностей при использовании вычислен-

ных на предыдущих шагах согласно (25) и (26) элементов корреляционной матрицы \mathbf{R} и вектора взаимной корреляции \mathbf{g} .

Погрешности, оценки которых представлены на рис. 3, обусловлены влиянием аддитивного шума наблюдения, однако на практике важное значение имеют также погрешности, возникающие из-за неточности вводимых шагов фазового сдвига.

На рис. 4а приведены значения погрешностей оценок фазы интерференционных полос, полученные в условиях случайных отклонений задаваемого фазового сдвига γ . Кривые получены для 20-ти реализаций случайных отклонений фазы по гауссову закону с нулевым средним и СКО, выраженным в долях от размера фазового сдвига. Из рисунка видно, что погрешности фазы не превышают сотых долей радиана при отклонениях фазового сдвига с СКО до 5%. Таким образом, рассма-

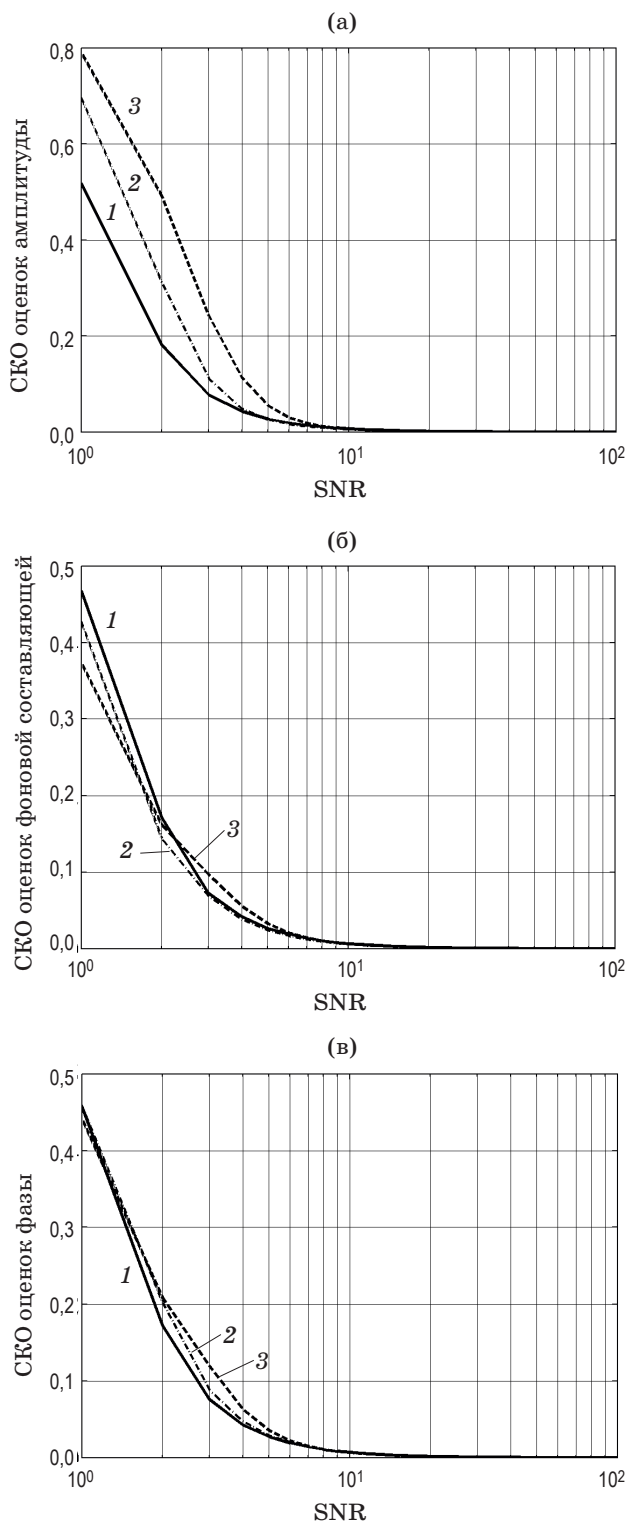


Рис. 3. Погрешности определения параметров двумерной картины интерференционных полос в зависимости от степени влияния аддитивного шума: погрешности оценок амплитуды (а), фоновой составляющей (б) и погрешности оценок фазы (в радианах) (в) при фазовых сдвигах $2\pi/4$ (1), $2\pi/7$ (2) и $2\pi/10$ (3).

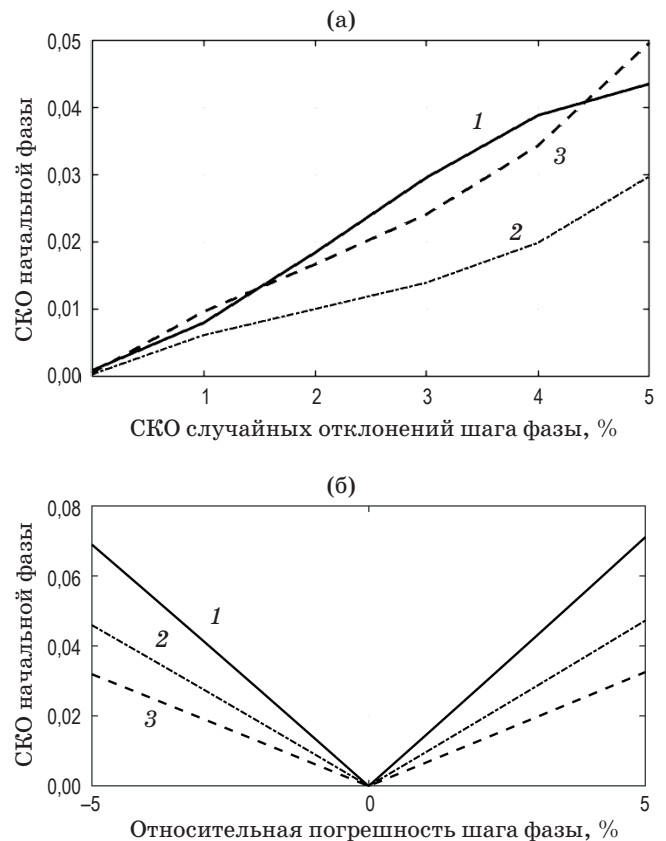


Рис. 4. Погрешности оценок фазы (в радианах) в зависимости от случайных отклонений шага фазового сдвига (а) и неточности калибровки фазового сдвига (б) при номинальных значениях фазового сдвига $2\pi/4$ (1), $2\pi/7$ (2) и $2\pi/10$ (3).

триваемый метод обеспечивает получение оценок фазы интерференционных полос с погрешностью, соответствующей наиболее точным алгоритмам в ИФС, погрешности которых приведены, например, в публикациях [1, 2, 4, 5].

Помимо случайных отклонений фазового сдвига необходимо учитывать также погрешность калибровки шага γ , поскольку это значение входит в формулу (22) для расчета фазы.

На рис. 4б представлены кривые, характеризующие систематическую погрешность оценок фазы по формуле (22) в случае неточной калибровки шага фазового сдвига γ . Для уменьшения систематических погрешностей требуется определять фактический размер фазового сдвига в конкретном эксперименте. Для решения этой задачи в предлагаемом методе используется методика, рассмотренная в последующем разделе статьи.

ОПРЕДЕЛЕНИЕ ФАКТИЧЕСКОГО РАЗМЕРА ФАЗОВОГО СДВИГА

Из выражения (22) видно, что для вычисления фазы интерференционных полос требуется знание фазового сдвига γ . На практике это значение может быть постоянным, но априорно не известным с высокой точностью. При этом используются методы обобщенной ИФС для точного определения фактического размера фазового сдвига [7, 8]. Применение рассматриваемого адаптивного ФВ открывает новые возможности решения этой задачи.

Очевидно, что сигнал в любом пикселе (i, j) по набору зарегистрированных интерференционных картин может быть принят в качестве опорного сигнала вместо априорно заданного в выражении (19) опорного сигнала. В частности, сигнал в пикселе (i, j) может быть принят как опорный по отношению к сигналу в соседнем пикселе. Ввиду близкого расположения соседних пикселей можно считать видность и фоновую составляющую интерференционных полос одинаковыми в этих пикселях. При этом выражения (21) и (23) упрощаются и принимают вид, соответственно

$$w_1^2 + w_2^2 + w_3^2 + 2(w_1w_2 + w_2w_3) \cos \gamma + 2w_1w_3 \cos 2\gamma = 1, \quad (30)$$

$$w_1 + w_2 + w_3 = 1. \quad (31)$$

Выражение (30) можно записать в форме квадратного уравнения

$$\alpha \cos^2 \gamma + \beta \cos \gamma + \rho = 0, \quad (32)$$

где $\alpha = 4w_1w_3$, $\beta = 2(w_1w_2 + w_2w_3)$, $\rho = (w_1 - w_3)^2 + w_2^2 - 1$, отсюда получаем оценку фактического фазового сдвига в форме

$$\hat{\gamma} = \arccos \left[\alpha^{-1} \left(-\beta/2 \pm \sqrt{(\beta/2)^2 - \alpha\rho} \right) \right]. \quad (33)$$

Два корня уравнения (33) возникают при возведении в квадрат выражения (21) (при $A = a$), вследствие чего отрицательные значения амплитуды (не имеющие физического смысла) также удовлетворяют уравнению (32). Априорное знание диапазона, в котором находится задаваемый фазовый сдвиг (например, от $2\pi/10$ до $2\pi/4$), позволяет выбрать правильную оценку шага $\hat{\gamma}$ в выражении (33).

Исследования показали, что корректная оценка соответствует выбору наибольшего из двух корней уравнения (33) в каждом пикселе (i, j) .

Алгоритм определения фактического фазового сдвига включает выполнение следующих процедур.

1. Осуществляется регистрация нескольких интерференционных картин с произвольным размером фазового сдвига.

2. Вычисляются весовые коэффициенты \hat{w} согласно выражению (8). Отметим, что вычисление элементов корреляционной матрицы \mathbf{R} и коэффициента корреляции g в (8) в соответствии с формулами (25) и (26) не требует знания фазового сдвига γ .

3. Из выражения (33) вычисляется фактический размер фазового сдвига.

Отметим, что дискриминант $(\beta/2)^2 - \alpha\rho$ в выражении (33) всегда неотрицателен (это можно показать, используя выражения (22), (30)–(32)). Однако на практике в условиях воздействия шума и при вычислениях с ограниченной точностью рекомендуется проверка знака дискриминанта с отбрасыванием пикселей, в которых оценки, получаемые согласно (33), не соответствуют допустимому диапазону значений шага фазового сдвига. Данная процедура не приводит к снижению точности ввиду того, что фазовый сдвиг одинаков для всех пикселей (i, j) , и получаемых корректных оценок достаточно для точного определения фактического значения фазового сдвига. Для дополнительного повышения точности целесообразно выполнить усреднение оценок по отобраным корректным значениям в пикселях (i, j) .

На рис. 5а, б в качестве примера представлены картина интерференционных полос и соответствующее ей распределение фазы. При обработке набора из 5-ти интерференционных картин с фазовыми сдвигами $2\pi/4$ вычислены весовые коэффициенты ФВ и дискриминант уравнения (32). Полученные значения дискриминанта для различных значений фазы иллюстрируются на рис. 5в. Подстановка значений дискриминанта и других коэффициентов (также зависящих от фазы интерференционных полос) в выражение (33) дает точную оценку фазового сдвига для всех пикселей (i, j) , как это показано на рис. 5г. Однако в условиях влияния шума и случайных отклонений шага фазового сдвига получаемые оценки отличаются от аналитически точных

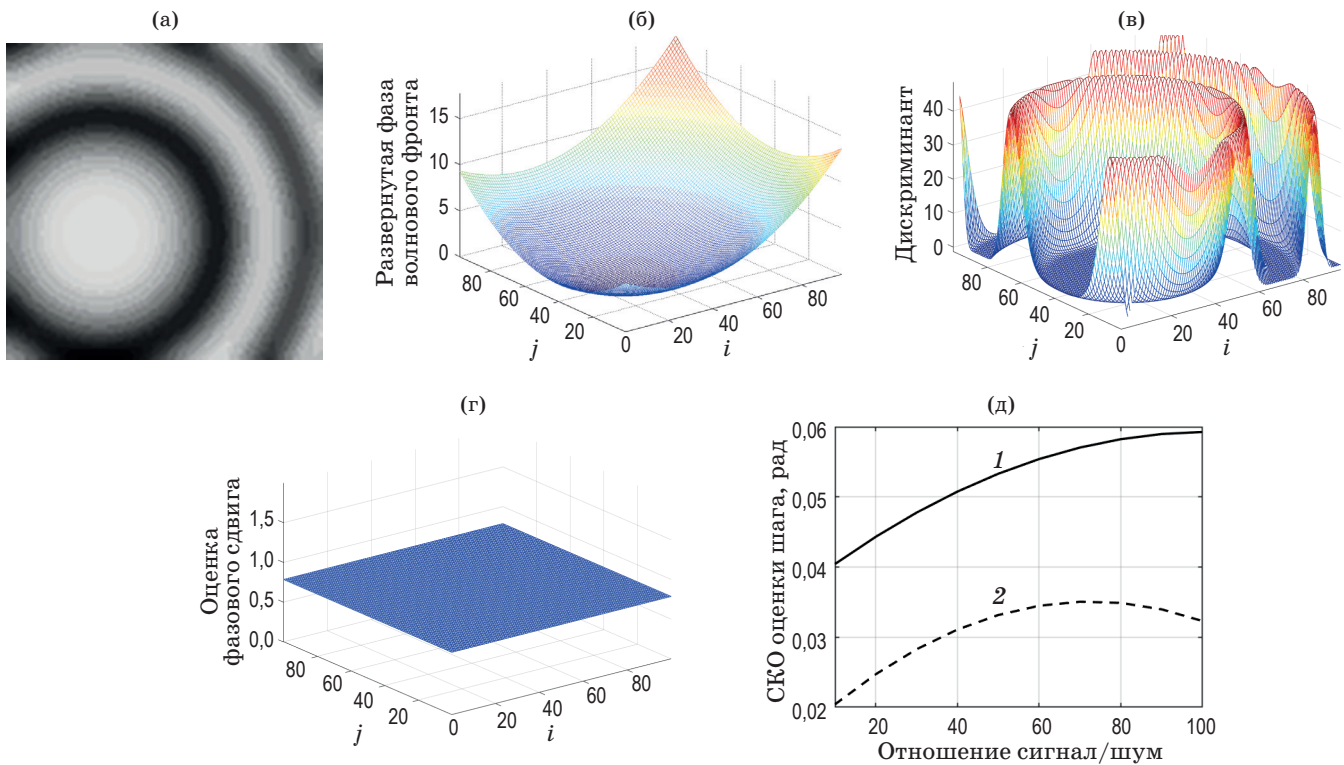


Рис. 5. Оценивание фазового сдвига: исходная интерференционная картина (а), развернутая фаза интерференционных полос (б), дискриминант уравнения (32) при $\gamma = 2\pi/4$ (в), оценка фактического фазового сдвига (г) и погрешности оценки фазового сдвига в условиях влияния помех при номинальных фазовых сдвигах $2\pi/4$ (1) и $2\pi/10$ (2) (д).

оценок. На рис. 5д представлены погрешности определения фазового сдвига в зависимости от отношения сигнал/шум (в определении (28)) при одновременном влиянии случайных отклонений фазового сдвига с СКО 3% от номинального значения шага. Кривые получены при усреднении данных по 20-ти реализациям. Характер кривых иллюстрирует особенность ФВ, заключающуюся в том, что наличие заметного аддитивного шума, не коррелированного с сигналом, снижает влияние случайных отклонений фазового сдвига за счет декорреляции помех по отношению к сигналу. Поэтому снижение уровня аддитивного шума приводит вначале к повышению погрешностей, которые затем имеют тенденцию к снижению при увеличении отношения сигнала к шуму.

ОБСУЖДЕНИЕ

В настоящей работе впервые показана возможность непосредственного вычисления параметров интерференционных полос через полученные коэффициенты адаптивного ФВ на

основе приведенных выше аналитических выражений, которые не требуют задания начальных условий.

Проведенное рассмотрение показывает, что ФВ позволяет определять параметры интерференционных полос с высокой точностью в динамическом режиме при малом количестве регистрируемых интерференционных картин. Приведенные выше оценки СКО фазы интерференционных полос на уровне $2\pi/100$ (и менее) соответствуют весьма малой погрешности $\lambda/200$ (λ — длина волны источника излучения) при определении формы волнового фронта.

Однако следует учитывать, что в основе рассматриваемого метода используется линейное уравнение Винера–Хопфа (7), непосредственное решение которого не предполагает получение оценок параметров, связанных нелинейно со значениями сигнала, в том числе фазы сигнала. Несмотря на то, что решение получено, и ФВ позволяет уменьшить влияние помех даже в том случае, когда их спектр накладывается на спектр сигнала (как это показано, например в [16]), применение ФВ наиболее

эффективно в случае влияния аддитивного шума, не коррелированного с сигналом. Поэтому полученные оценки погрешностей достигаются при сравнительно малых случайных отклонениях фазового сдвига (до нескольких процентов), когда в соответствии с выражением (27) возможна достаточно точная компенсация квадрата опорного сигнала суперпозицией взвешенных значений вектора взаимной корреляции входного и опорного сигналов.

Следует отметить, что точность метода возрастает при адекватном вычислении математического ожидания элементов корреляционной матрицы R сигнала и вектора взаимной корреляции g опорного и входного сигналов. Иными словами, отсутствие начальных условий компенсируется информацией, извлекаемой из корреляционных свойств сигналов. Погрешности алгоритма снижаются примерно как корень квадратный из общего количества последовательно обработанных интерференционных картин Q в выражениях (25) и (26).

В случае, если количество регистрируемых и обрабатываемых интерференционных картин критично, для динамического восстановления фазы целесообразно воспользоваться алгоритмом нелинейной калмановской фильтрации (см. [15]), в котором достаточно зарегистрировать три и обработать две интерференционные картины для получения развернутой фазы волнового фронта. Однако для этого необходимо наличие априорной информации об исследуемом процессе и задание начальных условий.

ЗАКЛЮЧЕНИЕ

Рассмотренный метод ИФС, основанный на использовании адаптивного ФВ, обеспечивает динамическое оценивание параметров интер-

ферометрических сигналов с высокой точностью. Метод можно отнести к классу обобщенных методов ИФС ввиду того, что он оптимизирован по критерию минимума дисперсии ошибки, позволяет использовать произвольное количество фазовых сдвигов и определять фактическое значение фазового сдвига.

Метод обладает высокой помехоустойчивостью по отношению к аддитивному шуму наблюдения благодаря суммированию регистрируемых отсчетов сигнала перед их обработкой и использованию корреляционного метода для вычисления весовых коэффициентов фильтра. Однако значительные случайные отклонения фазового сдвига (более 10%) могут приводить к заметному снижению точности ввиду того, что модель интерферометрического сигнала и выражения для оценивания его параметров выведены без непосредственного учета помех этого вида, и такое рассмотрение представляет задачу дальнейших исследований. Важное преимущество рассмотренного метода состоит в возможности задания опорного сигнала, содержащего адекватную информацию об исследуемом процессе.

Настраиваемые коэффициенты ФВ зависят от входного сигнала, поэтому рассматриваемый метод относится к нелинейной фильтрации. Представленные выше результаты получены для ФВ с тремя коэффициентами, т.е. с минимальным количеством коэффициентов, необходимых для определения трех неизвестных параметров: амплитуды, фоновой составляющей и фазы интерференционных полос.

Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации (проект №8.2501.2017/4.6).

ЛИТЕРАТУРА

1. Malacara D., Servin M., Malacara Z. Interferogram analysis for optical testing. NY: Taylor & Fransis, 2005. 568 p.
2. Schwider J., Burow R., Elssner K.-E., et al. Digital wavefront measuring interferometry: Some systematic error sources // Appl. Opt. 1983. V. 22. P. 3421–3432.
3. De Groot P.J. Vibration in phase-shifting interferometry // JOSA A. 1995. V. 12. P. 354–365.
4. Hariharan P., Oreb B.F., Eiju T. Digital phase-shifting interferometry: A simple error-compensating phase calculation algorithm // Appl. Opt. 1987. V. 26. P. 2504–2506.
5. Schmit J., Creath K. Extended averaging technique for derivation of error-compensating algorithms in phase-shifting interferometry // Appl. Opt. 1995. V. 34. P. 3610–3619.
6. Zhang H., Lalor M.J., Burton D.R. Error-compensating algorithms in phase-shifting interferometry: A comparison by error analysis // Opt. Lasers Eng. 1999. V. 31. P. 381–400.

7. *Patil A., Rastogi P.* Approaches in generalized phase shifting interferometry // *Opt. Lasers Eng.* 2005. V. 43. P. 475–490.
8. *Novák J., Novák P., Mikš A.* Multi-step phase-shifting algorithms insensitive to linear phase shift errors // *Opt. Commun.* 2008. V. 281. P. 5302–5309.
9. *Freischlad K., Koliopoulos C.L.* Fourier description of digital phase-measuring interferometry // *JOSA A.* 1990. V. 7. P. 542–551.
10. *Servin M., Estrada J.C.* Analysis and synthesis of phase shifting algorithms based on linear systems theory // *Opt. Lasers Eng.* 2012. V. 50. P. 1009–1014.
11. *Buytaert J.N., Dirckx J.J.J.* Study of the performance of 84 phase-shifting algorithms for interferometry // *J. Opt.* 2011. V. 40. P. 114–131.
12. *Wiener N.* Extrapolation, interpolation and smoothing of stationary time series. MIT Press, 1949. 163 p.
13. *Kalman R.E.* A new approach to linear filtering and prediction problems // *Trans. ASME.* 1960. V. 82. P. 35–45.
14. *Gurov I., Volynsky M.* Interference fringe analysis based on recurrence computational algorithms // *Opt. Lasers Eng.* 2012. V. 50. P. 514–521.
15. *Garifullin A., Gurov I., Volynsky M.* Unwrapped wavefront evaluation in phase-shifting interferometry based on 3D dynamic fringe processing in state space // *JOSA A.* 2016. V. 33. P. 1612–1621.
16. *Vaseghi S.V.* Advanced digital signal processing and noise reduction. 3rd ed. Chichester: John Wiley & Sons Ltd, 2006. Chapt. 6. P. 165–186.
17. *Уидроу Б., Стирнз С.* Адаптивная обработка сигналов. М.: Радио и связь, 1989. 440 с.
18. *Nava-Vega A., Salas L., Luna E., et al.* Correlation algorithm to recover the phase of a test surface using phase-shifting interferometry // *Opt. Exp.* 2004. V. 12. P. 5296–5306.