

Основные ограничения при синтезе цифровых голограмм 3D-изображений

*А. Л. Пазоев¹**

¹ Сибирский государственный университет геосистем и технологий, г. Новосибирск,
Российская Федерация

* e-mail: pazoev-al2018@sgugit.ru

Аннотация. Обсуждаются проблемы цифрового синтеза голограмм 3D объектов. Численно показана проблема использования быстрого преобразования Фурье (БПФ) при синтезе удаленного объекта. Предложен метод осуществления прямого численного преобразования Френеля, основанный на использовании пространственных симметрий ядра преобразования Френеля, как основного элемента вычислений. Обсуждаются его перспективные возможности. Приводится пример расчета голограммы 3D изображения и его восстановление. Разработана программа синтеза голограмм прямым интегральным преобразованием без использования БПФ, с помощью которой время синтеза голограмм сокращается более чем в четыре раза. Показаны перспективы этого метода, позволяющие сокращать время прямого синтеза голограмм 3D объектов на порядок и более.

Ключевые слова: голография, голографическая информация, дискретное преобразование Френеля, 3D объект

Main limitations in the synthesis of digital holograms of 3D images

*A. L. Pazoev¹**

¹ Siberian State University of Geosystems and Technologies, Novosibirsk, Russian Federation

* e-mail: pazoev-al2018@sgugit.ru

Abstract. Problems of digital synthesis of holograms of 3D objects are discussed. The problem of using fast Fourier transform (FFT) in the synthesis of a remote object is numerically shown. A method for implementing a direct numerical Fresnel transformation based on the use of spatial symmetries of the Fresnel transformation kernel as the main element of calculations is offered. Its promising possibilities are discussed. An example of calculating a hologram of a 3D image and its restoration is given. At the moment, a program for the synthesis of holograms by direct integral transformation without the use of FFT has been developed, with the help of which the time for the synthesis of holograms is reduced more than four times. The prospects of this method are shown, which make it possible to reduce the time of direct synthesis of holograms of 3D objects by an order of magnitude or more.

Keywords: holography, holographic information, discrete Fresnel transformation, 3D object

Введение

Одна из важнейших проблем численного синтеза и восстановления голограмм порождается их большой информационной емкостью и большими объемами вычислений интегральных преобразований, описывающих распространение света в пространстве. Особенную трудность представляет собой численный синтез голограмм 3D объекта.

Методы и модели

Без ограничения общности этот процесс можно представить на схеме Лейта-Упатниекса (рис. 1) [1], в которой объектный пучок формируется объектом 4 $U_0(x_0, y_0)$, находящимся на расстоянии z от плоскости регистрации голограммы (x_1, y_1) . В этой плоскости комплексная амплитуда электромагнитного светового поля $U_1(x_1, y_1)$, описываемая преобразованием Френеля, складывается с полем опорной волны $R_3(x_1, y_1)$. Интерференционное поле, формируемое в плоскости 5, регистрируется фотоприемным устройством, которым может служить как фотографический материал, так и матрица фотоприемников, и поэтому регистрируется уже не амплитуда, а распределение интенсивности $|U_1(x_1, y_1)|^2$ по полю голограммы.

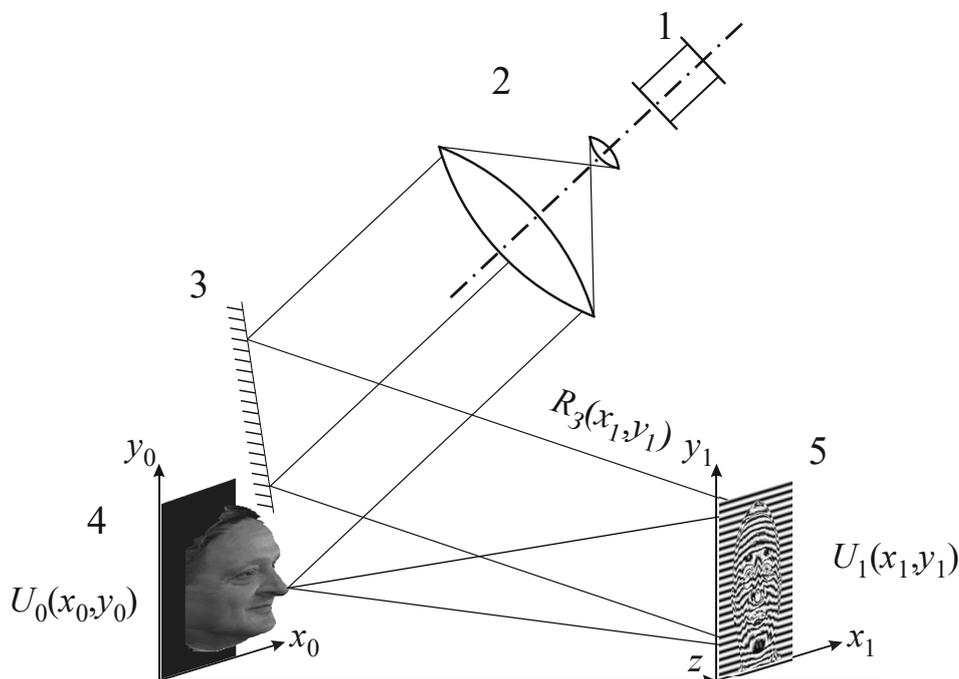


Рис. 1. Обобщенная схема записи голограмм Лейта и Упатниекса:

1 – лазер, 2 – коллимирующая система, 3 – зеркало, 4 – объект голографирования, 5 – фотоматериал, на котором регистрируется интерференционная картина, образуя голограмму

При освещении голограммы восстанавливающей волной, в простейшем случае равной опорной, при которой записывалась эта голограмма, восстанавливается 3 расходящихся порядка дифракции (рис. 2). Вблизи самой голограммы все 3 порядка наложены друг на друга, по мере отдаления порядки расходятся. Центральный (нулевой) порядок – это прошедший непроdifрагировавший на голограмме свет, минус первый порядок является восстановленным голограммой изображением объекта голографирования, а плюс первый – сопряженным изображением.

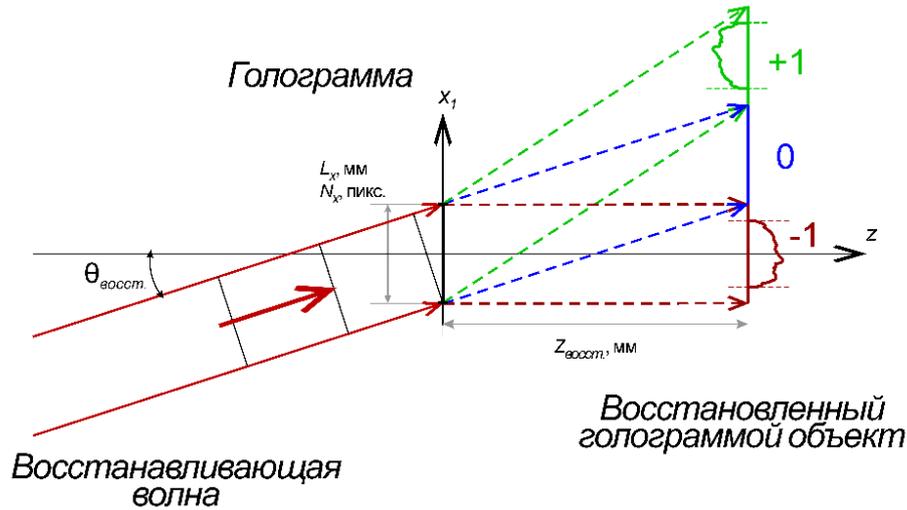


Рис. 2. Схема восстановления голограмм Лейта и Упатниекса, показаны расходящиеся порядки

Распространение излучения на рис. 1 в объектном пучке от объекта голографирования до голограммы, в однородной среде описывается преобразованием Френеля (1):

$$U_1(x_1; y_1) = \frac{e^{i \cdot k_\lambda \cdot z}}{i \cdot \lambda \cdot z} \iint U_0(x_0; y_0) \cdot \theta_{Fr.}(x_0; y_0; x_1; y_1) dx_0 dy_0, \quad (1)$$

где $U_0(x_0; y_0)$ – комплексная амплитуда излучения, сформированного объектом; $U_1(x_1; y_1)$ – комплексная амплитуда образа Френеля; λ – длина волны излучения; $k_\lambda = 2\pi/\lambda$ – волновое число; z – расстояние от объекта до фотоматериала, регистрирующего голограмму; $\theta_{Fr.}(x_0; y_0; x_1; y_1)$ – ядро интегрального преобразования, осуществляющее сложную в математическом смысле операцию – отображение всех точек объекта в каждую точку голограммы.

$$\theta_{Fr.}(x_0; y_0; x_1; y_1) = e^{\frac{i \cdot k_\lambda}{2 \cdot z} \cdot ((x_1 - x_0)^2 + (y_1 - y_0)^2)}. \quad (2)$$

Выражения (1–2) представлены в аналитическом виде, для их численного представления интегралы заменяются довольно громоздкими суммами (3):

$$U_1(x_{1j}; y_{1k}) = \sum_{p=1}^{Nz_p} \frac{e^{i \cdot k_\lambda \cdot z_p}}{i \cdot \lambda \cdot z_p} \sum_{n=1}^{Nx_{0n}} \sum_{m=1}^{Ny_{0m}} \sum_{j=1}^{Nx_{1j}} \sum_{k=1}^{Ny_{1k}} U_0(x_{0n}; y_{0m}) \cdot \theta_{Fr.}(x_{0n}; y_{0m}; x_{1j}; y_{1k}), \quad (3)$$

$$\theta_{Fr.}(x_{0n}; y_{0m}; x_{1j}; y_{1k}) = e^{\frac{i \cdot k_\lambda}{2 \cdot z_p} \cdot ((x_{1j} - x_{0n})^2 + (y_{1k} - y_{0m})^2)}, \quad (4)$$

где $U_0(x_{0n}; y_{0m})$ – комплексная амплитуда излучения, сформированного объектом; $U_1(x_{1j}; y_{1k})$ – комплексная амплитуда образа Френеля; $\theta_{Fr.}(x_{0n}; y_{0m}; x_{1j}; y_{1k})$ – ядро преобразования; λ – длина волны излучения; $k_\lambda = 2\pi/\lambda$ – волновое число; z_p – расстояние от объекта до фотоматериала, регистрирующего голограмму; Nx_{0n}, Ny_{0m}, Nz_p – количество отсчетов, представляющих объект голографирования, соответственно по осям x, y, z ; Nx_{1j}, Ny_{1k} – количество отсчетов по x, y на голограмме.

Видно, что объемы прямых вычислений велики. Действительно, при прямом вычислении преобразования по формулам (3, 4) операцию суммирования формально надо выполнять $(N_{x0} \times N_{y0}) \times (N_{x1} \times N_{y1}) \times Nz_p$ раз, т. е. при равенстве указанных переменных – N^5 раз.

При больших размерах, при условии, что реальный размер пикселя должен удовлетворять критерию Найквиста [2], который следует из теоремы отсчетов Котельникова [3], когда ширина интерференционной полосы должна содержать больше чем два отсчета. При периоде интерференционных полос, задаваемых углом падения опорной волны, порядка $\lambda/\sin(\theta/2)$ и расстоянии до голограммы, равном ее размеру ($\theta \sim 1$ рад), ширина интерференционной полосы будет около ~ 1 мкм. В этом случае для синтеза голограммы размером А4 (210×297 мм) количество вычисляемых значений суммы (3) будет равно числу точек объекта, умноженному на число вычисляемых точек в голограмме. Для объекта размером с кадр высокоразрешающего современного TV изображения, как Full HD, это $1920 \times 1080 \approx 2,1 \cdot 10^6$ и голограммы, размером А4 (210×297 мм² = $6,2 \cdot 10^{10}$ мкм²) число представляемых в дискретном виде точек голограммы и при двух пикселях на точку, размером 1 мкм, составляет примерно $1,2 \cdot 10^{11}$. Последовательное вычисление ядер (4), а затем суммирование такого количества элементов в (3) даже для кодирования каждой точки двумя пикселями, потребует более 10^{17} операций. Мы использовали не специализированный компьютер с процессором (модель AMD Athlon X4 860K) с тактовой частотой 3,7 ГГц. При этом вычисление одного ядра (4) составляло примерно 32 нс, без учета дополнительных задержек в циклах чтения/записи информации на физический носитель. Легко оценить, что прямое вычисление преобразования Френеля только лишь для одного полного слоя по z для указанных реальных размеров объекта и голограммы заняло бы более 8×10^9 с, т. е. 255 лет, что исключает возможность прямых расчетов по (3).

Обычно [4] эту проблему решают использованием быстрого преобразования Фурье (БПФ), когда преобразование Френеля вычисляется с его использованием, как обратное БПФ от произведения спектра объекта на передаточную функцию (5, 6), алгоритм D-FFT (Double Fast Fourier Transform):

$$U_1(x_{1j}; y_{1k}) = FFT^{-1} \left\{ FFT \{ U_0(x_{0n}; y_{0m}) \} \cdot H_F(f_{x1j}; f_{y1k}) \right\}, \quad (5)$$

$$H_F(f_{x1j}; f_{y1k}) = \exp \left[ik_\lambda z_p \sqrt{1 - \lambda^2 (f_{x1j}^2 - f_{y1k}^2)} \right], \quad (6)$$

где $FFT\{U_0(x_{0n}; y_{0m})\}$ – спектр объекта; $H_F(f_{x1j}; f_{y1k})$ – передаточная функция; f_{x1j}, f_{y1k} – частотные координаты.

Такой подход существенно сокращает вычисления, однако он тоже не совершенен и в ряде задач просто не может быть применен. Например, при вычислениях БПФ теряется фаза при переходе через 2π , т. е. вычисления Фурье образа объектов реальных размеров прямыми вычислениями часто представляет затруднение. Для снятия этого противоречия разрабатывались методы развертки фазы, например [5–7], которые тоже ограничивают скорость вычислений. Однако это не единственная проблема. Дело в том, что БПФ требует равенства представления в пикселях объекта голографирования и самой голограммы. Это сразу увеличивает количество вычислений, поскольку сравнительно малый размер (в пикселях) объекта (в нашем случае 2×10^6) сразу заменяется на 10^{11} и число суммирований в (5) возрастает в 10^5 раз, существенным образом нивелируя полученный с применением БПФ выигрыш.

Еще одно требование к голограммам 3D объектов заключается в необходимости обеспечить удобство наблюдения их параллакса. Для этого необходимо, чтобы наблюдаемое поле 3D сцены находилось в области проекции взгляда наблюдателя на голограмму, не выходя за ее края (рис. 3). Это тоже увеличивает поле голограммы. Для наблюдения одного горизонтального параллакса в угле $\pm 45^\circ$, надо увеличивать размер голограммы примерно в 3 раза по сравнению с размером объекта голографирования.

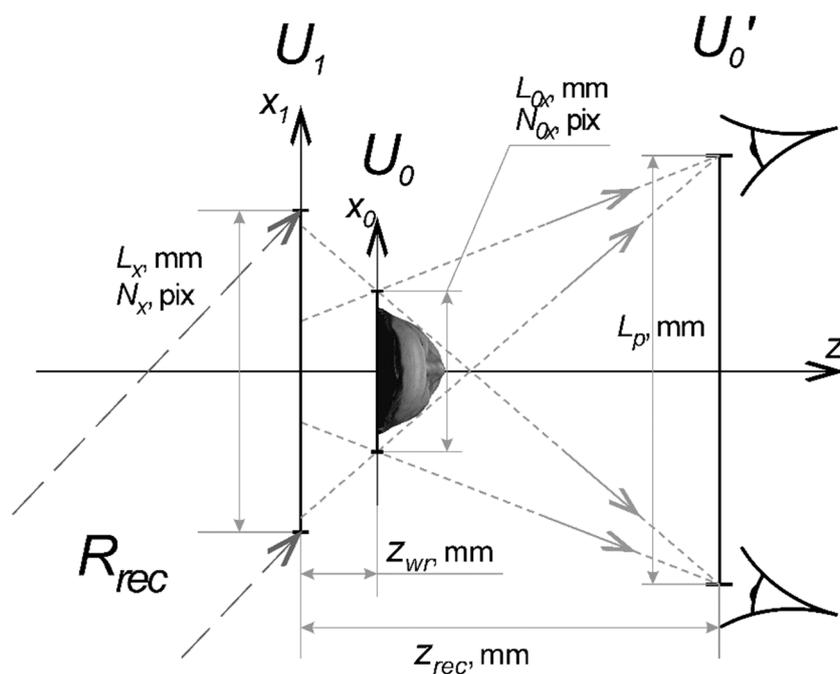


Рис. 3. Условие наблюдения параллакса

Особенно такое увеличение вычислений мешает синтезу голограмм 3D объекта, поскольку БПФ, лежащее в основе D-FFT (5), работает только с плоскими изображениями. При этом, единственным способом вычисления D-FFT является разбиение изображения по слоям. На рис. 4 показан пример разбиения 3D изображения по слоям: 4а на 16 слоев, 4б на 256 слоев.

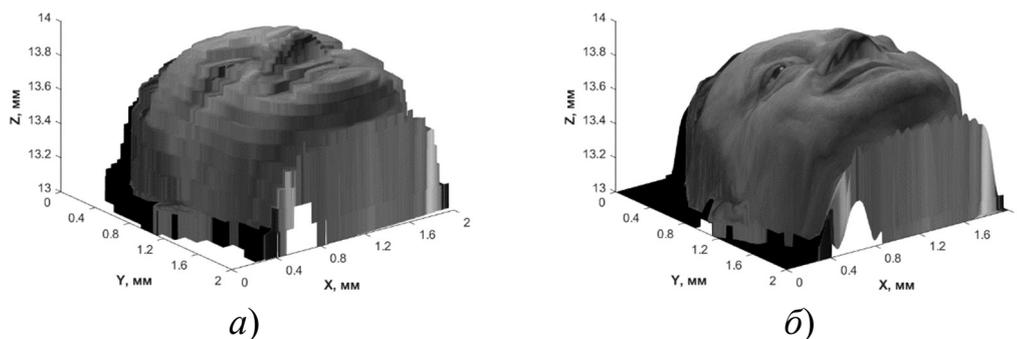


Рис. 4. 3D объект, разрезанный на слои вдоль оси z:
 а) на 16 слоев; б) на 256 слоев

Видно, что при 256 слоях объект гладкий, а при 16 слоях явно видны ступеньки. Количество вычислений в (5) увеличивается пропорционально числу слоев. И если при прямом вычислении возможно отсутствующие в очередном слое точки исключить из расчетов, то при использовании D-FFT, в силу особенностей БПФ, этого сделать нельзя. БПФ и основанное на нем D-FFT могут считать только полные плоскости, несмотря на то, что значительная часть вычислений будет от пустых (нулевых) значений исходной функции $U_0(x_0; y_0)$. При разбиении 3D объекта на 256 слоев, что требуется для ощущения непрерывного параллакса, увеличение времени расчетов возрастает пропорционально, существенно гася выигрыш, даваемый использованием D-FFT. Так, произведенная численная оценка на примере простой геометрической фигуры – 3D конуса, при его разрезании на 256 слоев, содержит 99,69 % нулевых пикселей.

В связи с этим возникает задача поиска других, не использующих БПФ способов и вычисления дискретного преобразования Френеля.

Одним из них является метод использования свойств симметрий ядер преобразования Френеля [8–10]. В этом случае можно оптимизировать вычисления, во-первых, исключая избыточность вычислений путем использования свойств симметрий в ядрах преобразования, а во-вторых, пропуская пустые вычисления на нулевых пикселях.

Результаты

На основе предложенного метода разрабатывается программа на базе MATLAB, в которой используются некоторые, хотя пока еще и не все возможные, симметрии в ядрах преобразования Френеля (рис. 5).

Была проведена оценка времени вычисления дискретного преобразования Френеля (ДПФр) предложенным методом, для голограммы размером 2000×2000 точек от 3D объекта размером 1000×1000 точек, время вычисления ДПФр указанным методом составит ~38 часов, что более чем в 4 раза быстрее, чем если бы вычисление проходило прямым суммированием по (3, 4).

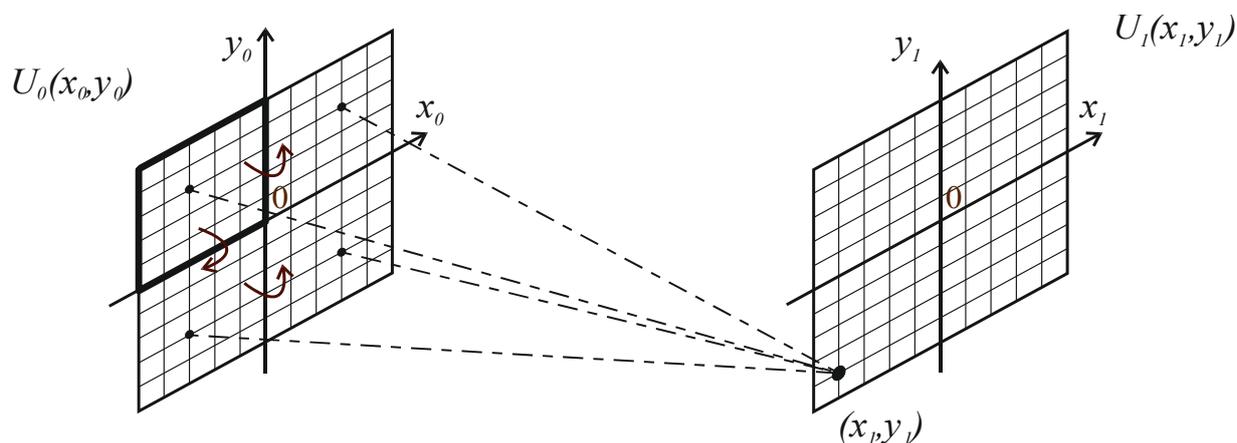


Рис. 5. Осевые симметрии в ядрах преобразования Френеля. Плоскость $U_0(x_0, y_0)$ – объект, плоскость $U_1(x_1, y_1)$ – голограмма

При использовании алгоритма D-FFT вычисление образа Френеля для голограммы размером 2000×2000 точек от 3D объекта размером уже 2000×2000 точек (как было упомянуто ранее, из-за особенностей БПФ размер объекта должен быть равен размеру голограммы), при разделении его на 16 слоев по глубине (z) время вычисления составило 4 мин.

Видно, что реализация предложенного метода на основе симметрий в ядрах ДПФр пока недостаточно совершенна, хотя уже и существенно лучше, чем прямое вычисление ДПФр [8–10]. Однако потенциальные возможности велики. Это можно оценить (рис. 6) подсчетом числа уникальных ядер в преобразованиях Френеля и Фурье, на основе которого построен метод D-FFT.

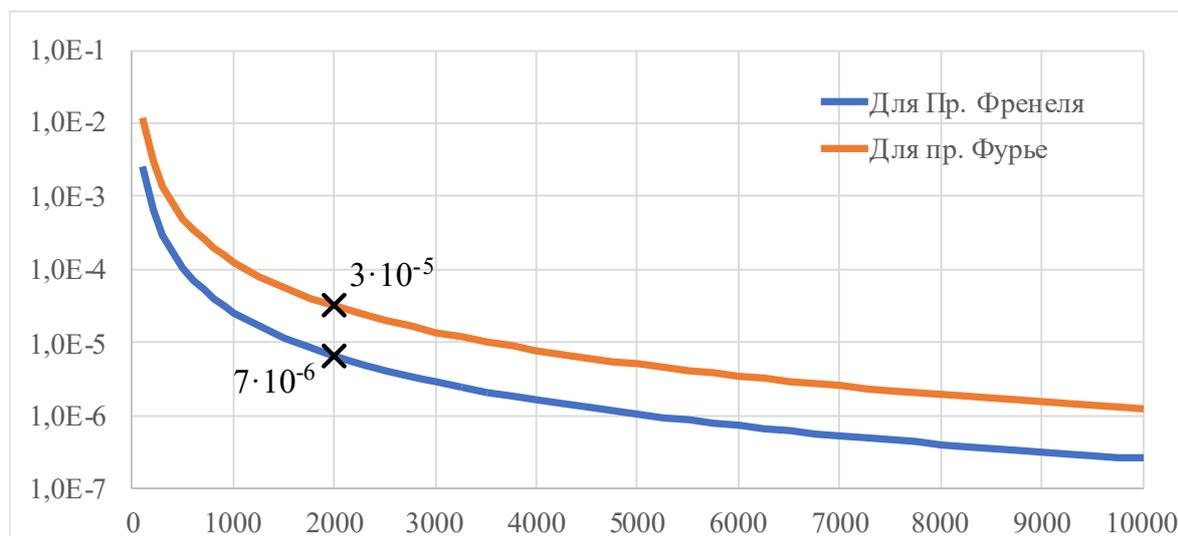


Рис. 6. Зависимость процентного отношения количества уникальных ядер к общему их количеству (N^4) от линейного размера объекта (N): оранжевым для преобразования Фурье, синим для преобразования Френеля. Точки на графиках соответствуют размеру голограммы рис. 7

Видно, что процентное отношение количества уникальных ядер к их общему количеству в преобразовании Френеля на порядок меньше чем в преобразовании Фурье. Если учесть все обозначенные в работе недостатки метода D-FFT, основанного на быстром преобразовании Фурье, то можно сделать вывод, что предлагаемый метод на основе симметрий в ядрах преобразования Френеля потенциально более перспективен, для чего требуется тщательнее определить и затем использовать в (3) все возможные степени симметрии ядра преобразования Френеля (4).

Тем не менее, ускорение в расчетах, полученное с использованием 4 видов симметрии, уже позволяет проводить расчеты голограмм 3D объектов в реальном времени. На рис. 7 приведена одна из них.

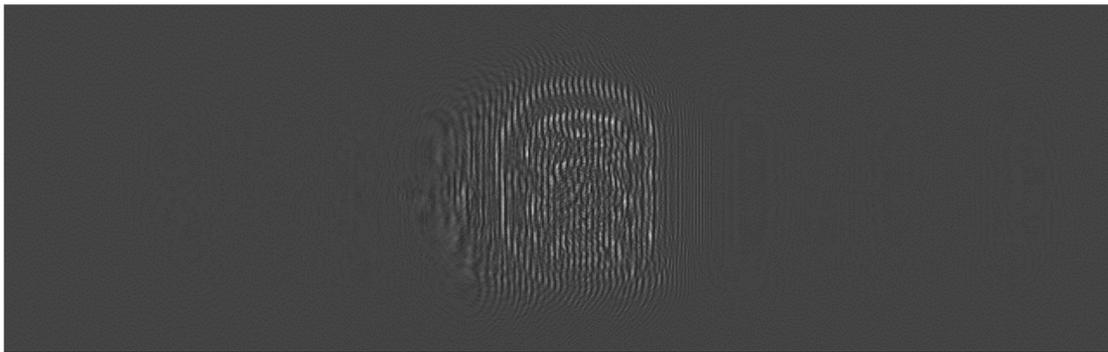


Рис. 7. Голограмма, синтезированная от 3D объекта, разрезанного на 16 слоев (рис. 4а)

На рис. 7 показана голограмма, размером 2000×2000 пикс., синтезированная от 3D объекта размером 2000×2000 пикс., разрезанного на 16 слоев (рис. 4а).

Восстановленное изображение, также полученное предложенным методом преобразования Френеля, показано на рис. 8. Очевидно, что при увеличении числа слоев до 256 восстановленное изображение может восприниматься как изображение с непрерывным параллаксом.

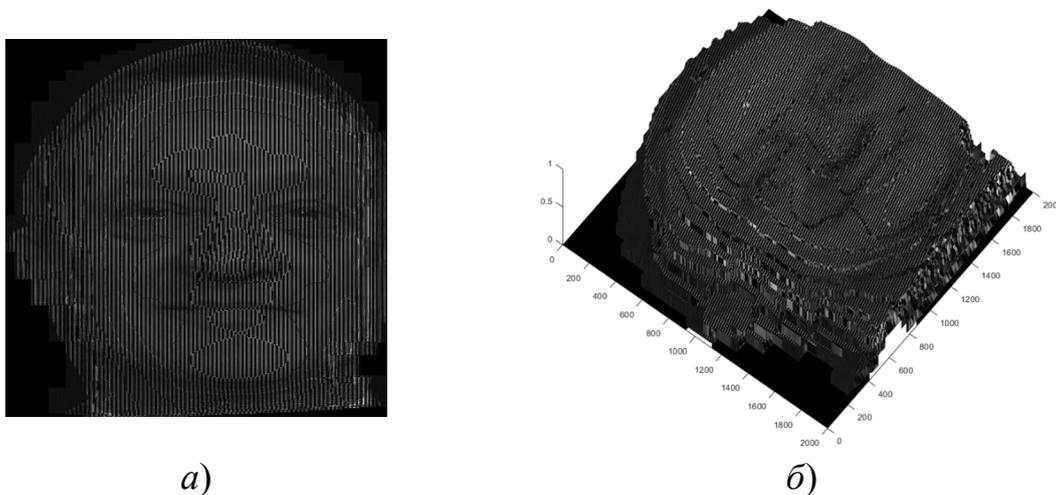


Рис. 8. Восстановленный голограммой (рис. 7) 3D объект

Заключение

Показаны основные проблемы синтеза голограмм 3D объектов численными методами, связанные с ограничениями, присущими БПФ. Предлагается альтернативный метод вычисления преобразования Френеля без использования БПФ. Приводится пример численного синтеза голограммы по известной численной модели 3D голографируемого объекта и восстановления записанного на ней 3D изображения. Обсуждаются перспективы предлагаемого метода вычислений.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Leith E. N. Reconstructed Wavefronts and Communication Theory / E. N. Leith, J. Upatnieks // Journal of the Optical Society of America. – 1962. – vol. 52. – P.1123-1130.
2. Nyquist H. Certain factors affecting telegraph speed / H. Nyquist // The Bell System Technical Journal. – 1924. – Т. 3. – № 2. – С. 324-346. <https://doi.org/10.1002/j.1538-7305.1924.tb01361.x>.
3. Котельников В. А. О пропускной способности эфира и проволоки в электросвязи – Всесоюзный энергетический комитет. // Материалы к I Всесоюзному съезду по вопросам технической реконструкции дела связи и развития слаботочной промышленности, 1933. Репринт статьи в журнале УФН. – Т. 176. – № 7. – 2006. – С. 762-770.
4. Pascal Picart, Jun-chang Li. Digital Holography. John Wiley & Sons, 2012, P. 88-92.
5. Шевкунов И. А. Метод фазовой развертки сдвигом восстановленного поля / И. А. Шевкунов // Вестник Санкт-Петербургского университета. – 2015. – Т. 2. – № 60. – С. 395-401.
6. Yan L. A robust phase unwrapping algorithm based on reliability mask and weighted minimum least-squares method / L. Yan, H. Zhang, R. Zhang, X. Xie, B. Chen // Optics and Lasers in Engineering. – 2019. – Т. 112. – С. 39-45.
7. Xie X. Iterated unscented Kalman filter for phase unwrapping of interferometric fringes / X. Xie // Optics Express. – 2016. – Т. 24. – № 17. – С. 18872-18897.
8. Пазоев А. Л. Преобразование Френеля в моделировании голограмм / А. Л. Пазоев, С. А. Шойдин // Сборник материалов XIV Международного научного конгресса и выставки ИНТЕРЭКСПО ГЕО-СИБИРЬ-2018. – 2018. – № 9. – С. 95-98.
9. Пазоев А. Л. Быстрое преобразование Френеля в среде Матлаб / А. Л. Пазоев, С. А. Шойдин // Сборник материалов XIV Международного научного конгресса и выставки ИНТЕРЭКСПО ГЕО-СИБИРЬ-2018. – 2018. – № 9. – С. 90-94.
10. Шойдин С. А. Использование пространственных симметрий для задач моделирования процессов передачи изображений / С. А. Шойдин, А. Л. Пазоев // Тезисы докладов XXV Международного Симпозиума Оптика атмосферы и океана. Физика атмосферы. – 2019. – С. 90-94.

© А. Л. Пазоев, 2022