

Рис. 1: Оптическая схема

УДК 519.688

Использование некоторых оптимизационных методов в задачах когерентной оптики и голографии

А.А. Распопов

Московский государственный университет им. М. В. Ломоносова

1. Введение.

Голография — специальный метод построения изображений предметов. Условно этот процесс можно разделить на два этапа. На первом предмет освещают пучком света, волновое поле рассеивается предметным, и опорная волна записывается на специальном экране — голограмме. На втором этапе голограмму освещают опорной волной, формируя изображение предмета. Рассмотрим более подробно, как это происходит.

Пусть в плоскости $z = 0$ (рис.1) находится объект S с заданным распределением электромагнитного поля $E(x,y)$. На расстоянии d от объектной плоскости расположена голографическая пластинка Γ . Запишем распределение объектного поля \tilde{E} в плоскости голограммы Γ :

$$\tilde{E}(\xi, \eta) = \iint_S K(\xi - x, \eta - y, d) E(x, y) dx dy$$

где $K(x, y, z) = -\frac{\partial}{\partial z} \left(\frac{e^{ikr}}{2\pi r} \right)$, $r = \sqrt{x^2 + y^2 + z^2}$, $k = 2\pi/\lambda$, λ - длина волны. перепишем, используя преобразование Фурье:

$$\tilde{E}(\xi, \eta) = \frac{1}{2\pi} \left[\iint_{\xi^2 + \eta^2 < k^2} e^{i(-x\xi - y\eta + d\sqrt{k^2 - \xi^2 - \eta^2})} H(\xi, \eta) d\xi d\eta + \iint_{\xi^2 + \eta^2 > k^2} e^{i(-x\xi - y\eta) - d\sqrt{\xi^2 + \eta^2 - k^2}} H(\xi, \eta) d\xi d\eta \right]$$

здесь $H(\xi, \eta)$ определяется соотношением

$$E(\xi, \eta) = \frac{1}{2\pi} \iint_S e^{-i(x\xi + y\eta)} H(\xi, \eta) d\xi d\eta$$

Таким образом, мы получили поле, совпадающее с предметным, т.е. расходящееся от оригинала. Чтобы получить поле, сходящееся к оригиналу, т.е. инверсное к предметному полю, надо в качестве

значений поля на голограмме взять функцию, комплексно сопряженную к $\tilde{E}(x,y)$. В результате восстановления с голограммы в плоскости объекта возникнет близкое, но, тем не менее, отличное от $E(x,y)$, поле:

$$E_1(x,y) = \frac{1}{2\pi} \left[\iint_{\xi^2+\eta^2 < k^2} e^{i(x\xi+y\eta)} H^*(\xi,\eta) d\xi d\eta + \iint_{\xi^2+\eta^2 > k^2} e^{i(x\xi+y\eta)-2d\sqrt{\xi^2+\eta^2-k^2}} H^*(\xi,\eta) d\xi d\eta \right]$$

Получается, посредством любой голограммы нельзя восстановить точное предметное изображение. Введём оператор размытия B , сопоставляющий исходному предметному полю E восстановленное поле E_1 , которое получается с помощью формул выше. Он задается следующей формулой:

$$B(E)(p,q) = \iint_S P(\xi,\eta,p,q) d\xi d\eta,$$

где

$$P(\xi,\eta,p,q) = \iint_{\Gamma} K(x-p,y-q,d)K^*(x-p,y-q,d) d\xi d\eta.$$

Функция $P(\xi,\eta,p,q)$ называется ядром размытия. Размытое изображение $B(E)$ – это результат композиции размытия отдельных точек исходного изображения. Можно сформулировать следующую задачу оптимизации: необходимо найти такое поле $G(x,y)$, что при его записи на голограмму и последующем восстановлении, полученное поле $B(G)(x,y)$ было максимально "близко" к заданному $E(x,y)$:

$$\min_G D\{E(x,y), B(G)(x,y)\},$$

где D - некоторый функционал, показывающий "близость" двух полей.

2. Описание алгоритма.

Для численного решения поставленной задачи выполним дискретизацию. Разобьем топологии на ячейки одинакового размера S_{ij} . Тогда полям (x,y) и $G(x,y)$ будут соответствовать дискретные функции: $E(x_i,y_j)$ и $G(x_i,y_j)$ для всех $i = 1, \dots, n, j = 1, \dots, m$, где (x_i,y_j) - координаты центра ячейки S_{ij} . Таким образом, можно сопоставить $G(x,y)$ и $E(x,y)$ матрицы \tilde{G} и \tilde{E} . Оператор размытия представляет собой дискретную свертку, задаваемую формулой $B(\tilde{G}) = \tilde{G} * \tilde{P}$, где \tilde{P} - комплекснозначная матрица, соответствующая ядру оператора размытия. Задача минимизации становится следующей:

$$\min_G D\{\tilde{E}(x,y), B(\tilde{G})(x,y)\}$$

Перед дальнейшим рассказом давайте введем следующие обозначения. Так как после дискретизации фактически вся работа ведется с матрицами, основным параметром будет пара (i,j) - положение в соответствующей комплекснозначной матрице.

$X(i,j) = \{x_{i,j} + iy_{i,j}\}$ - предсказанное изображение.

$\bar{X}(i,j) = \{\bar{x}_{i,j}\}$ - идеальное изображение, комплексной части нет, причем $\bar{x}_{i,j} \in \{0,1\}$.

$P(i,j) = \{p_{i,j} + iq_{i,j}\}$ - оператор размытия.

$B(i,j) = \{ReB(i,j) + iImB(i,j)\} = \{\sum_{s,t} P(i-s,j-t)X(s,t)\} = \{\sum_{s,t} (p_{i-s,j-t}x_{s,t} - q_{i-s,j-t}y_{s,t}) + i(p_{i-s,j-t}y_{s,t} + q_{i-s,j-t}x_{s,t})\}$ - восстановленное изображение.

$I(i,j) = \{|B(i,j)|^2\} = \{(\sum_{s,t} (p_{i-s,j-t}x_{s,t} - q_{i-s,j-t}y_{s,t}))^2 + (\sum_{s,t} (p_{i-s,j-t}y_{s,t} + q_{i-s,j-t}x_{s,t}))^2\}$ - интенсивность восстановленного изображения.

$\bar{I}(i, j) = \{|\bar{x}(i, j)|^2\} = \{\bar{x}_{i,j}\}$ - интенсивность идеального изображения.

Дискретные градиенты матрицы интенсивностей:

$$\nabla_i I(i, j) = \frac{I(i+1, j) - I(i-1, j)}{2}, \nabla_j I(i, j) = \frac{I(i, j+1) - I(i, j-1)}{2}, \nabla I(i, j) = \sqrt{[\nabla_i I(i, j)]^2 + [\nabla_j I(i, j)]^2}.$$

Также нам понадобится весовая функция $W(i, j)$, которая задается следующим образом:

$$W(i, j) = [\bar{I}(i, j) \oplus S(i, j)] - [\bar{I}(i, j) \ominus S(i, j)]$$

$S(i, j)$ - матрица 3x3 из одних единиц, а операции \oplus и \ominus называются расширение и сужение соответственно и формально задаются следующим образом:

$$A \ominus B = \{z | (B)_z \subseteq A\}$$

$$A \oplus B = \{z | [(B)_z \cap A] \subseteq A\}$$

Сужение A по B - это все такие элементы z , если z - центр B , то $B \subseteq A$. Расширение A по B - это все такие элементы z , если z - центр B , то B пересекается с A .

Таким образом функция $W(i, j)$ "вырезает" границу поля. Рассмотрим подробнее, как это происходит в нашей задаче. Нам нужны сужение и расширение $\bar{I}(i, j) \oplus S(i, j)$, $\bar{I}(i, j) \ominus S(i, j)$. $S(i, j)$ - матрица $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, а $\bar{I}(i, j)$ - тоже матрица, на самом деле состоящая из нулей и единиц. В случае, например, когда $\bar{I}(i, j)$ - матрица 9x9

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

соответствующие расширение и сужение будут следующими:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Перейдем к описанию функционала D . Будем рассматривать

$$D\{\bar{I}(i, j), I(i, j)\} = L\{\bar{I}(i, j), I(i, j)\} + \gamma_1 R_{TV}\{I(i, j)\} + \gamma_2 R_{contrast}\{\bar{I}(i, j), I(i, j)\}$$

Первое слагаемое - расстояние по метрике L_2

$$L\{\bar{I}(i, j), I(i, j)\} = \sum_{i, j} (I(i, j) - \bar{I}(i, j))^2$$

Второе слагаемое функционала, оператор полной вариации, определим как

$$R_{TV}\{I(i, j)\} = \sum_{i, j} (1 - W(i, j)) \cdot \nabla I(i, j),$$

и третье слагаемое, оператор контрастности:

$$R_{contrast}\{I(i, j), \bar{I}(i, j)\} = \sum_{i, j} W(i, j) \cdot \{|\nabla_i [I(i, j) - \bar{I}(i, j)]| + |\nabla_j [I(i, j) - \bar{I}(i, j)]|\}.$$

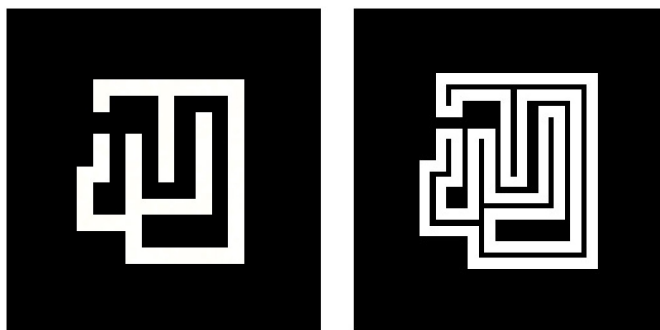


Рис. 2: Идеальное изображение и его граница.

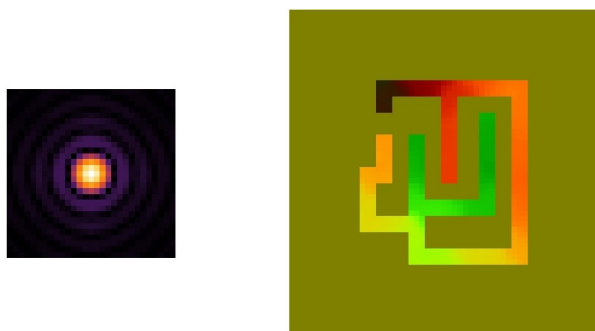


Рис. 3: Пятно размытия и стартовая позиция

Почему мы выбрали именно такой функционал, и как он поможет приблизить восстановленное изображение к желаемому? Во-первых, в задачах голографии важно не само изображение, а его интенсивности $I(i, j)$ в каждой точке, а первое слагаемое как раз некоторым образом измеряет различие в интенсивностях двух изображений. Во-вторых, граница изображения - самое уязвимое место при размытии, тогда как нужно, чтобы она была максимально четкой. Третье слагаемое функционала позволяет сделать границу наиболее похожей на границу идеального изображения. На участках вне границы соседние точки должны мало отличаться, что и обеспечивает второе слагаемое.

Перед нами стоит задача минимизировать $D\{\bar{I}(i, j), I(i, j)\}$ по переменным $x_{i,j}, y_{i,j}$ матрицы предсказанного изображения. Решать задачу будем методом градиентного спуска по формуле $I = I - \alpha \cdot \text{grad}D\{\bar{I}(i, j), I(i, j)\}$. Вычислим необходимые градиенты.

$$\begin{aligned} L\{\bar{I}(i, j), I(i, j)\} &= \sum_{i,j} (I(i, j) - \bar{I}(i, j))^2 = \\ &= \sum_{i,j} \left(\left[\sum_{s,t} (p_{i-s,j-t}x_{s,t} - q_{i-s,j-t}y_{s,t}) \right]^2 + \left[\sum_{s,t} (p_{i-s,j-t}y_{s,t} + q_{i-s,j-t}x_{s,t}) \right]^2 - \bar{x}_{i,j} \right)^2; \\ \frac{\partial L}{\partial x_{k,l}} &= 4 \sum_{i,j} (I(i, j) - \bar{I}(i, j)) \cdot [p_{i-k,j-l} \text{Re}B(i, j) + q_{i-k,j-l} \text{Im}B(i, j)]; \\ \frac{\partial L}{\partial y_{k,l}} &= 4 \sum_{i,j} (I(i, j) - \bar{I}(i, j)) \cdot [(-q_{i-k,j-l}) \text{Re}B(i, j) + p_{i-k,j-l} \text{Im}B(i, j)]. \end{aligned}$$

Градиент оператора полной вариации:

$$\begin{aligned} R_{TV}\{I(i, j)\} &= \sum_{i,j} (1-W(i, j)) \cdot \nabla I(i, j) = \sum_{i,j} (1-W(i, j)) \cdot \left(\left| \frac{1}{2}(I(i+1, j) - I(i-1, j)) \right|^2 + \left| \frac{1}{2}(I(i, j+1) - I(i, j-1)) \right|^2 \right)^{\frac{1}{2}} = \\ &= \frac{1}{2} \sum_{i,j} (1-W(i, j)) \cdot \left[\left[\sum_{s,t} (p_{i+1-s,j-t}x_{s,t} - q_{i+1-s,j-t}y_{s,t}) \right]^2 + \left[\sum_{s,t} (p_{i+1-s,j-t}y_{s,t} + q_{i+1-s,j-t}x_{s,t}) \right]^2 - \right. \\ &\quad \left. - \left[\sum_{s,t} (p_{i-1-s,j-t}x_{s,t} - q_{i-1-s,j-t}y_{s,t}) \right]^2 - \left[\sum_{s,t} (p_{i-1-s,j-t}y_{s,t} + q_{i-1-s,j-t}x_{s,t}) \right]^2 + \right. \\ &\quad \left. + \left[\sum_{s,t} (p_{i-s,j+1-t}x_{s,t} - q_{i-s,j+1-t}y_{s,t}) \right]^2 + \left[\sum_{s,t} (p_{i-s,j+1-t}y_{s,t} + q_{i-s,j+1-t}x_{s,t}) \right]^2 - \right. \\ &\quad \left. - \left[\sum_{s,t} (p_{i-s,j-1-t}x_{s,t} - q_{i-s,j-1-t}y_{s,t}) \right]^2 + \left[\sum_{s,t} (p_{i-s,j-1-t}y_{s,t} + q_{i-s,j-1-t}x_{s,t}) \right]^2 \right]^{\frac{1}{2}}; \\ \frac{\partial R_{TV}}{\partial x_{k,l}} &= \sum_{i,j} (1-W(i, j)) \cdot \frac{1}{\nabla I(i, j)} \cdot [\nabla_i I(i, j)(p_{i+1-k,j-l} \text{Re}B(i+1, j) + q_{i+1-k,j-l} \text{Im}B(i+1, j)) - \\ &\quad - p_{i-1-k,j-l} \text{Re}B(i-1, j) - q_{i-1-k,j-l} \text{Im}B(i-1, j)) + \nabla_j I(i, j)(p_{i-k,j+1-l} \text{Re}B(i, j+1) + q_{i-k,j+1-l} \text{Im}B(i, j+1)) - \\ &\quad - p_{i-k,j-1-l} \text{Re}B(i, j-1) - q_{i-k,j-1-l} \text{Im}B(i, j-1)]; \\ \frac{\partial R_{TV}}{\partial y_{k,l}} &= \sum_{i,j} (1-W(i, j)) \cdot \frac{1}{\nabla I(i, j)} \cdot [\nabla_i I(i, j)(-q_{i+1-k,j-l} \text{Re}B(i+1, j) + p_{i+1-k,j-l} \text{Im}B(i+1, j)) + \\ &\quad + q_{i-1-k,j-l} \text{Re}B(i-1, j) - p_{i-1-k,j-l} \text{Im}B(i-1, j)) + \nabla_j I(i, j)(-q_{i-k,j+1-l} \text{Re}B(i, j+1) + p_{i-k,j+1-l} \text{Im}B(i, j+1)) + \\ &\quad + q_{i-k,j-1-l} \text{Re}B(i, j-1) - p_{i-k,j-1-l} \text{Im}B(i, j-1)]; \end{aligned}$$

$$+q_{i-k,j-1-l}ReB(i,j-1) - p_{i-k,j-1-l}ImB(i,j-1)].$$

И наконец, градиент оператора контрастности.

$$\begin{aligned} R_{contrast}\{I(i,j), \bar{I}(i,j)\} &= \sum_{i,j} W(i,j) \cdot \{|\nabla_i [I(i,j) - \bar{I}(i,j)]| + |\nabla_j [I(i,j) - \bar{I}(i,j)]|\} = \\ &= \frac{1}{2} \sum_{i,j} W(i,j) \cdot (|\sum_{s,t} (p_{i+1-s,j-t}x_{s,t} - q_{i+1-s,j-t}y_{s,t})|^2 + |\sum_{s,t} (p_{i+1-s,j-t}y_{s,t} + q_{i+1-s,j-t}x_{s,t})|^2 - \\ &- [\sum_{s,t} (p_{i-1-s,j-t}x_{s,t} - q_{i-1-s,j-t}y_{s,t})]^2 - [\sum_{s,t} (p_{i-1-s,j-t}y_{s,t} + q_{i-1-s,j-t}x_{s,t})]^2 + \bar{x}_{i+1,j} - \bar{x}_{i-1,j}| + \\ &+ |[\sum_{s,t} (p_{i-s,j+1-t}x_{s,t} - q_{i-s,j+1-t}y_{s,t})]^2 + [\sum_{s,t} (p_{i-s,j+1-t}y_{s,t} + q_{i-s,j+1-t}x_{s,t})]^2 - \\ &- [\sum_{s,t} (p_{i-s,j-1-t}x_{s,t} - q_{i-s,j-1-t}y_{s,t})]^2 + [\sum_{s,t} (p_{i-s,j-1-t}y_{s,t} + q_{i-s,j-1-t}x_{s,t})]^2 + \bar{x}_{i,j+1} - \bar{x}_{i,j-1}|); \end{aligned}$$

$$\begin{aligned} \frac{\partial R_{contrast}}{\partial x_{k,l}} &= \sum_{i,j} W(i,j) \cdot (sgn(\nabla_i [I(i,j) - \bar{I}(i,j)]) \cdot (p_{i+1-k,j-l}ReB(i+1,j) + q_{i+1-k,j-l}ImB(i+1,j)) - \\ &- p_{i-1-k,j-l}ReB(i-1,j) - q_{i-1-k,j-l}ImB(i-1,j)) + sgn(\nabla_j [I(i,j) - \bar{I}(i,j)]) \cdot (p_{i-k,j+1-l}ReB(i,j+1) + \\ &+ q_{i-k,j+1-l}ImB(i,j+1) - p_{i-k,j-1-l}ReB(i,j-1) - q_{i-k,j-1-l}ImB(i,j-1))); \end{aligned}$$

$$\begin{aligned} \frac{\partial R_{contrast}}{\partial y_{k,l}} &= \sum_{i,j} W(i,j) \cdot (sgn(\nabla_i [I(i,j) - \bar{I}(i,j)]) \cdot (-q_{i+1-k,j-l}ReB(i+1,j) + p_{i+1-k,j-l}ImB(i+1,j)) + \\ &+ q_{i-1-k,j-l}ReB(i-1,j) - p_{i-1-k,j-l}ImB(i-1,j)) + sgn(\nabla_j [I(i,j) - \bar{I}(i,j)]) \cdot (-q_{i-k,j+1-l}ReB(i,j+1) + \\ &+ p_{i-k,j+1-l}ImB(i,j+1) + q_{i-k,j-1-l}ReB(i,j-1) - p_{i-k,j-1-l}ImB(i,j-1))). \end{aligned}$$

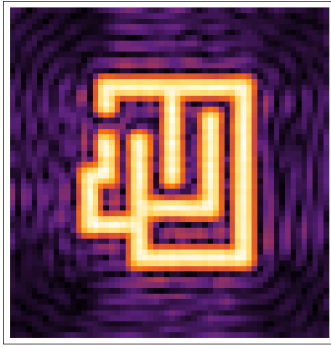


Рис. 4: $\gamma_1 = \gamma_2 = 0$

3. Результаты расчетов.

Будем решать задачу для изображения размером 60×58 дискретных точек, приведенного на рисунке 2. Пятно размытия 29×29 , в качестве стартовой позиции для градиентного спуска возьмем фазированное идеальное изображение, рис.3. При $\gamma_1 = \gamma_2 = 0$ задача была разобрана достаточно подробно ранее. Наиболее успешный результат приведен на рисунке 4. Видно, как размывается граница и появляются помехи на участках вне основного изображения.

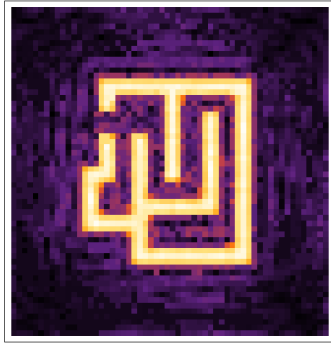


Рис. 5: $\gamma_1 = 0.33333333, \gamma_2 = 0.1$

На рисунке выше приведен результат работы с полным функционалом для значений $\gamma_1 = 0.33333333, \gamma_2 = 0.1$. При этих значениях все три слагаемых в случае исследуемого идеального изображения принимают почти одинаковые значения. Заметно, что по сравнению с рис.4, намного улучшилась граница. В обоих случаях было сделано $T=1000$ итераций с шагом $L=0.0001$ и точностью $\varepsilon = 0.0000001$. Выбранные значения γ_1 и γ_2 , судя по всему являются близкими к оптимальным. Также следует упомянуть, что количество итераций имеет существенное значение. На рис.6 приведены результаты для $T=100$ итераций с различными параметрами γ_1, γ_2 .

Таким образом, цель работы достигнута, получено достаточно хорошее предсказание заданного изображения. Построенный функционал выполняет возложенные на него задачи — улучшает границу.

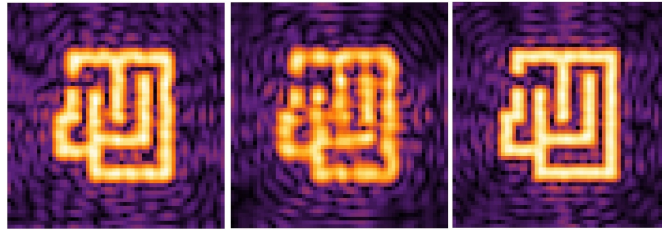


Рис. 6: Результат для следующих параметров: $\gamma_1 = 0.7, \gamma_2 = 0.4$; $\gamma_1 = 1, \gamma_2 = 1$; $\gamma_1 = 0.3333333, \gamma_2 = 0.1$

Список литературы

- [1] БОРИСОВ М.В., БОРОВИКОВ В.А., ГАВРИКОВ А.А., КНЯЗЬКОВ Д.Ю., РАХОВСКИЙ В.И., ЧЕЛЮБЕЕВ Д.А., ШАМАЕВ А.С.: *Методы создания и коррекции качества голографических изображений геометрических объектов с элементами субволновых размеров*. Доклады Академии Наук. 2010. Т. 434, № 3. С. 332–336.
- [2] GABOR D. A.: *A New Microscopic Principle*. Nature, 1948.
- [3] NINGNING JIA AND EDMUND Y. LAM: *Pixelated source mask optimization for process robustness in optical lithography*. Optics express, 2011.
- [4] R. C. GONZALEZ AND R. E. WOODS: *Digital Image Processing, 2nd ed.*. Prentice Hall, 2002.