Градиентный метод расчета каскадных ДОЭ и его применение в задаче классификации рукописных цифр

Д.В. Сошников [1,2], Л.Л. Досколович [1,2], Е.В. Бызов [1,2]
[1] ИСОИ РАН – филиал ФНИЦ «Кристаллография и фотоника» РАН,
443001, Россия, г. Самара, ул. Молодогвардейская, д. 151,
[2] Самарский национальный исследовательский университет имени академика С.П. Королёва,
443086, Россия, г. Самара, Московское шоссе, д. 34

Аннотация

Рассмотрен градиентный метод расчета каскадных дифракционных оптических элементов, состоящих из нескольких последовательно расположенных фазовых дифракционных оптических элементов. С использованием свойства унитарности оператора распространения света через каскадный дифракционный оптический элемент получены явные выражения для производных функционала ошибки по фазовым функциям каскадного дифракционного оптического элемента. Рассмотрено применение градиентного метода в задаче фокусировки различных падающих пучков в области с различными распределениями интенсивности и в задаче классификации изображений. Представленные описания градиентного метода рассматривают задачи синтеза каскадных дифракционных оптических элементов для фокусировки лазерного излучения и для классификации изображений в рамках единого методологического подхода. При этом показано, что вычисление производных функционалов ошибок и в задаче фокусировки, и в задаче классификации сводится к одной и той же общей формуле. С использованием предложенного градиентного метода рассчитаны одиночные и каскадный дифракционный оптический элемент для решения задачи классификации рукописных цифр. Полученные результаты могут найти применение при разработке дифракционных нейронных сетей и систем для фокусировки лазерного излучения.

Ключевые слова: дифракционный оптический элемент, фазовая функция, скалярная теория дифракции, градиентный метод, классификация изображений.

Цитирование: **Сошников, Д.В.** Градиентный метод расчета каскадных ДОЭ и его применение в задаче классификации рукописных цифр / Д.В. Сошников, Л.Л. Досколович, Е.В. Бызов // Компьютерная оптика. – 2023. – Т. 47, № 5. – С. 691-701. – DOI: 10.18287/2412-6179-CO-1314.

Citation: Soshnikov DV, Doskolovich LL, Byzov EV. Gradient method for designing cascaded DOEs and its application in the problem of classifying handwritten digits. Computer Optics 2023; 47(5): 691-701. DOI: 10.18287/2412-6179-CO-1314.

Введение

В последние годы дифракционные оптические элементы (ДОЭ) вновь стали предметом интенсивных исследований [1 – 5]. Основными причинами интереса к данной области исследований являются компактность, технологичность и эффективность применения ДОЭ при решении широкого класса задач по преобразованию и фокусировке оптического излучения. Как правило, расчет ДОЭ осуществляется в рамках скалярной теории дифракции. Задача расчета ДОЭ принадлежит к классу некорректных обратных задач и состоит в определении формы «фазового» дифракционного микрорельефа, обеспечивающего формирование в заданной области пространства светового поля с заданными характеристиками (как правило, с требуемым распределением интенсивности). Поскольку высота микрорельефа ДОЭ пропорциональна фазовой функции светового поля, формируемого ДОЭ, то задачу расчета ДОЭ обычно рассматривают как задачу расчета фазовой функции, обеспечивающей формирование требуемого распределения интенсивности. Для расчета фазовой функции традиционно используются различные итерационные алгоритмы, включающие «классический» алгоритм Гершберга–Сакстона (англ. Gerchberg–Saxton algorithm), алгоритм уменьшения ошибки и широкий спектр их различных модификаций [6 – 12].

Помимо одиночных ДОЭ, широко применяются т.н. каскадные ДОЭ, состоящие из нескольких последовательно расположенных фазовых ДОЭ. Такие ДОЭ обладают более широкими функциональными возможностями и позволяют решать более сложные задачи, включающие фокусировку различных падающих пучков в различные области [2, 13, 14], в том числе при различных длинах волн падающего излучения. Для расчета каскадных ДОЭ также используют итерационные алгоритмы, являющиеся обобщением алгоритмов, предложенных для одиночных ДОЭ. В то же время большинство из этих алгоритмов являются эвристическими и не имеют строгого теоретического обоснования. В частности, для итерационных алгоритмов, использованных в работах [2, 13, 14], отсутствует анализ свойств неувеличения ошибки, кото-

рыми обладают алгоритм Гершберга–Сакстона и алгоритм уменьшения ошибки. В этой связи представляется актуальной разработка градиентных методов расчета каскадных ДОЭ, которые являются более понятными с теоретической точки зрения.

Помимо применения каскадных ДОЭ в сложных задачах фокусировки лазерного излучения, в последние годы они нашли широкое применение в решении различных задач машинного обучения (в особенности задач классификации изображений) [3, 5, 15–21], в задачах реализации различных математических преобразований, описываемых линейными операторами [22], а также при реализации операций спектральной фильтрации [23]. В данных задачах каскадные ДОЭ стали называть дифракционными нейронными сетями (ДНН). Основным методом расчета ДНН является стохастический градиентный метод, а также основанные на нем «улучшенные» методы 1-го порядка [24]. В ряде работ были получены выражения для градиентов функций ошибок по параметрам фазовых функций [2, 5, 18]. Тем не менее, данные выражения были получены при дискретной записи операторов распространения света между ДОЭ. Вследствие этого выражения для градиентов имеют сложный вид и, по мнению авторов настоящей статьи, сложны для понимания и реализации.

В настоящей работе рассмотрен градиентный метод расчета каскадного ДОЭ. С использованием свойства унитарности оператора распространения света через каскадный ДОЭ получены явные и компактные выражения для производных функционала ошибки по фазовым функциям каскадного ДОЭ. Рассмотрено применение градиентного метода в задаче фокусировки различных падающих пучков в области с различными распределениями интенсивности и в задаче классификации изображений. Представленные описания градиентного метода «объединяют» задачи синтеза каскадных ДОЭ для фокусировки лазерного излучения и для классификации изображений в рамках единого методологического подхода. При этом показано, что несмотря на различный вид функционалов ошибки, используемых в задачах фокусировки (формирования заданных распределений интенсивности) и в задачах оптической классификации, вычисление производных данных функционалов сводится к одной и той же общей формуле. С использованием предложенного градиентного метода рассчитаны одиночные и каскадный ДОЭ для решения задачи классификации рукописных цифр. Представленные результаты численного моделирования демонстрируют хорошие рабочие характеристики предложенного метода.

1. Постановка задачи

Пусть во входной плоскости $z=f_0=0$ задана комплексная амплитуда «входного» поля $w_0(\mathbf{u}_0)$, где $\mathbf{u}_0=(u_0, v_0)$ – декартовы координаты в плоскости $z=0$. Будем считать, что световое поле (с длиной волны λ) из плоскости $z=0$ последовательно распространяется через набор из n фазовых ДОЭ, расположенных в плоскостях $z=f_1, \ldots, z=f_n$ $(0<f_1<\ldots<f_n)$, и далее попадает в выходную плоскость $z=f_{n+1}>f_n$ (рис. 1).

Обозначим $\varphi_1(\mathbf{u}_1),\ldots,\varphi_n(\mathbf{u}_n)$ фазовые функции ДОЭ, где $\mathbf{u}_j=(u_j, v_j)$ – декартовы координаты в плоскостях $z=f_1, \ldots, z=f_n$. Будем предполагать, что распространение света между плоскостями $z=f_i, i=1\ldots, n+1$ описывается в рамках приближения Френеля–Кирхгофа скалярной теории дифракции. Будем также считать, что прохождение света через ДОЭ описывается в приближении тонкого оптического элемента как умножение комплексной амплитуды падающего пучка на функцию комплексного пропускания ДОЭ

$$T_m(\mathbf{u}_m)=\exp\left\{i\varphi_m(\mathbf{u}_m)\right\}, m=1,\ldots n.$$

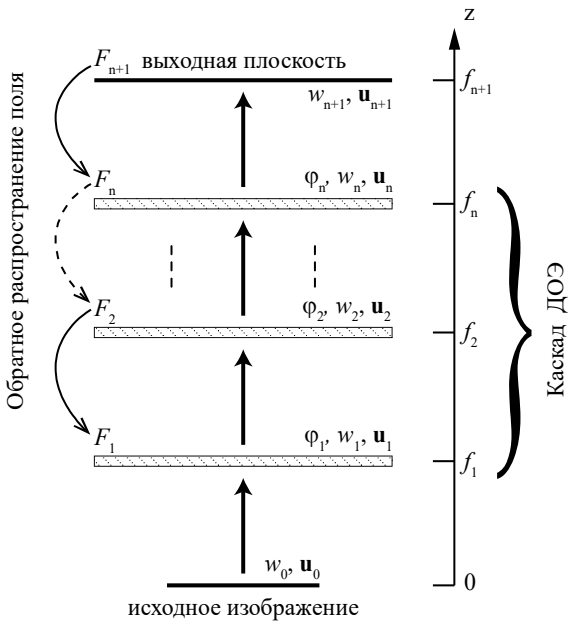


Рис. 1. Геометрия задачи расчета каскадного ДОЭ

В этом случае распространение света через каскадный ДОЭ описывается следующими формулами:

$$w_1(\mathbf{u}_1)=\frac{e^{ikd_1}}{\lambda i d_1}\iint w_0(\mathbf{u}_0)\exp\left\{i\frac{\pi}{\lambda d_1}(\mathbf{u}_1-\mathbf{u}_0)^2\right\}d^2\mathbf{u}_0,$$

$$w_m(\mathbf{u}_m)=\frac{e^{ikd_m}}{\lambda i d_m}\iint w_{m-1}(\mathbf{u}_{m-1})e^{i\varphi_{m-1}(\mathbf{u}_{m-1})}\times \quad (1)$$

$$\times\exp\left\{i\frac{\pi}{\lambda d_m}(\mathbf{u}_m-\mathbf{u}_{m-1})^2\right\}d^2\mathbf{u}_{m-1},\ m=2,\ldots,n+1,$$

где $d_m=f_m-f_{m-1}$ – расстояние между плоскостями. Согласно (1), вычисление комплексной амплитуды выходного поля $w_{n+1}(\mathbf{u}_{n+1})$ осуществляется рекуррентным образом. Для дальнейшего анализа формулы (1) будет удобно рассматривать как представление линейных операторов прямого распространения света из входной плоскости $z=f_0$ до плоскостей $z=f_m$, $m=1,\ldots,n+1$.

Под обратной задачей будем понимать задачу расчета фазовых функций $\varphi_1(\mathbf{u}_1),\ldots,\varphi_n(\mathbf{u}_n)$ из условия формирования в выходной плоскости светового поля с заданным распределением интенсивности $I(\mathbf{u}_{n+1})$. Будем считать, что ошибка формирования заданного распределения интенсивности представляется некоторым интегральным критерием

$$\varepsilon(\varphi_1,...,\varphi_n)=\iint D[I_{n+1}(\mathbf{u}_{n+1}),I(\mathbf{u}_{n+1})]d^2\mathbf{u}_{n+1}, \qquad (2)$$

где $I_{n+1}(\mathbf{u}_{n+1})=|w_{n+1}(\mathbf{u}_{n+1})|^2$ – распределение интенсивности, формируемое при фазовых функциях $\varphi_1(\mathbf{u}_1),\ldots,\varphi_n(\mathbf{u}_n)$, а D – некоторая функция, представляющая отличие формируемого и заданного распределений в текущей точке.

Далее мы будем рассматривать обратную задачу расчета каскадного ДОЭ как задачу минимизации функционала (2)

$$\varepsilon(\varphi_1,...,\varphi_n)\to\min_{\varphi_1,...,\varphi_n}. \qquad (3)$$

2. Градиентный метод расчета каскадного ДОЭ

Для решения задачи (3) будем использовать градиентный метод. Рассмотрим вычисление производной функционала (3) по функции φ_m.

Обозначим

$$\Delta_m\varepsilon(\varphi_1,...,\varphi_n)=\varepsilon(\varphi_1,...,\varphi_m+\Delta\varphi_m,...,\varphi_n)-\varepsilon(\varphi_1,...,\varphi_m,...,\varphi_n) \qquad (4)$$

приращение функционала ошибки, вызванное приращением $\Delta\varphi_m$ функции φ_m. Согласно (2), данное приращение имеет вид:

$$\begin{aligned}
\Delta_m\varepsilon(\varphi_1,...,\varphi_n)&=\iint\frac{\partial D[I_{n+1}(\mathbf{u}_{n+1}),I(\mathbf{u}_{n+1})]}{\partial I_{n+1}}\Delta_m\left(I_{n+1}(\mathbf{u}_{n+1})\right)d^2\mathbf{u}_{n+1}=\\
&=\iint\frac{\partial D[I_{n+1}(\mathbf{u}_{n+1}),I(\mathbf{u}_{n+1})]}{\partial I_{n+1}}\Delta_m\left(w_{n+1}(\mathbf{u}_{n+1})w^*_{n+1}(\mathbf{u}_{n+1})\right)d^2\mathbf{u}_{n+1}=\\
&=\iint\frac{\partial D[I_{n+1}(\mathbf{u}_{n+1}),I(\mathbf{u}_{n+1})]}{\partial I_{n+1}}2\,\mathrm{Re}\left[[\Delta_m w_{n+1}(\mathbf{u}_{n+1})]w^*_{n+1}(\mathbf{u}_{n+1})\right]d^2\mathbf{u}_{n+1}=\\
&=2\,\mathrm{Re}\left[\iint\Delta_m w_{n+1}(\mathbf{u}_{n+1})F^*_{n+1}(\mathbf{u}_{n+1})d^2\mathbf{u}_{n+1}\right]=2\,\mathrm{Re}\left\langle\Delta_m w_{n+1}(\mathbf{u}_{n+1}),F_{n+1}(\mathbf{u}_{n+1})\right\rangle,
\end{aligned} \qquad (5)$$

где $\Delta_m w_{n+1}(\mathbf{u}_{n+1})$ – приращение комплексной амплитуды, вызванное приращением фазы $\Delta\varphi_m$, угловые скобки обозначают скалярное произведение функций, а функция $F_{n+1}(\mathbf{u}_{n+1})$ имеет вид:

$$F_{n+1}(\mathbf{u}_{n+1})=\frac{\partial D[I_{n+1}(\mathbf{u}_{n+1}),I(\mathbf{u}_{n+1})]}{\partial I_{n+1}}w_{n+1}(\mathbf{u}_{n+1}). \qquad (6)$$

Для дальнейших выкладок введем оператор $\mathrm{Pr}_{f_{n+1}\to f_m^+}$ обратного распространения света из выходной плоскости $z=f_{n+1}$ до плоскости $z=f_m^+$, расположенной непосредственно за плоскостью расположения m-го ДОЭ $z=f_m$. В данной плоскости комплексная амплитуда поля при прямом распространении света имеет вид $w_m e^{i\varphi_m}$. Приведем формулы для вычисления данного оператора на примере поля $F_{n+1}(\mathbf{u}_{n+1})$ (рис. 1). При $m=n$

$$\begin{aligned}
F_n(\mathbf{u}_n)&=\mathrm{Pr}_{f_{n+1}\to f_n^+}(F_{n+1})=\frac{e^{-ikd_{n+1}}}{-\lambda i d_{n+1}}\times\\
&\times\iint F_{n+1}(\mathbf{u}_{n+1})\exp\left\{-i\frac{\pi}{\lambda d_{n+1}}(\mathbf{u}_n-\mathbf{u}_{n+1})^2\right\}d^2\mathbf{u}_{n+1}.
\end{aligned} \qquad (7)$$

Отметим, что выражение (7) является интегралом Френеля–Кирхгофа, в котором расстояние распространения d_{n+1} взято со знаком минус. При $m<n$ оператор $\mathrm{Pr}_{f_{n+1}\to f_m^+}(F_{n+1})$ рассчитывается рекуррентно на основе следующей формулы:

$$\begin{aligned}
F_{j-1}(\mathbf{u}_{j-1})&=\frac{e^{-ikd_j}}{-\lambda i d_j}\iint F_j(\mathbf{u}_j)e^{-i\varphi_j(\mathbf{u}_j)}\times\\
&\times\exp\left\{-i\frac{\pi}{\lambda d_j}(\mathbf{u}_{j-1}-\mathbf{u}_j)^2\right\}d^2\mathbf{u}_j,\ \ j=n,...,m+1.
\end{aligned} \qquad (8)$$

Несложно показать, что операторы прямого и обратного распространения света через набор фазовых ДОЭ являются унитарными и сохраняют скалярное произведение [16]. В силу сохранения скалярного произведения приращение критерия (5) можно представить в виде:

$$\begin{aligned}
\Delta_m\varepsilon(\varphi_1,...,\varphi_n)&=2\,\mathrm{Re}\left\langle\Delta_m w_{n+1},F_{n+1}\right\rangle=\\
&=2\,\mathrm{Re}\left\langle\mathrm{Pr}_{f_{n+1}\to f_m^+}(\Delta_m w_{n+1}),\mathrm{Pr}_{f_{n+1}\to f_m^+}(F_{n+1})\right\rangle.
\end{aligned} \qquad (9)$$

Поскольку $\mathrm{Pr}_{f_{n+1}\to f_m^+}(\Delta_m w_{n+1})=\Delta_m(w_m e^{i\varphi_m})$, где $w_m e^{i\varphi_m}$ – комплексная амплитуда поля непосредственно за плоскостью m-го ДОЭ при прямом распространении, и $\mathrm{Pr}_{f_{n+1}\to f_m^+}(F_{n+1})=F_m$, то преобразуем (9) к виду:

$$\begin{aligned}
\Delta_m\varepsilon(\varphi_1,...,\varphi_n)&=2\,\mathrm{Re}\left\langle\Delta_m(w_m e^{i\varphi_m}),F_m\right\rangle=\\
&=2\,\mathrm{Re}\iint w_m(\mathbf{u}_m)\Delta e^{i\varphi_m(\mathbf{u}_m)}F^*_m(\mathbf{u}_m)d^2\mathbf{u}_m.
\end{aligned} \qquad (10)$$

Представляя приращение $\Delta e^{i\varphi m}$ в виде разложения в ряд Тейлора до линейных членов

$$\Delta\exp(i\varphi m)=\exp(i\varphi m+i\Delta\varphi m)-\exp(i\varphi m)\approx i\Delta\varphi_m\exp(i\varphi m),$$

запишем главную линейную часть приращения функционала (10) в виде:

$$\delta_m\varepsilon\left(\varphi_1,...,\varphi_n\right)=-2\iint\Delta\varphi_m\left(\mathbf{u}_m\right)\times \\ \times\operatorname{Im}[w_m\left(\mathbf{u}_m\right)e^{i\varphi_m(\mathbf{u}_m)}F_m^*\left(\mathbf{u}_m\right)]d^2\mathbf{u}_m. \quad (11)$$

Согласно (11), производная функционала имеет вид:

$$\frac{\delta\varepsilon\left(\varphi_1,...,\varphi_n\right)}{\delta\varphi_m}=-2\operatorname{Im}[w_m\left(\mathbf{u}_m\right)e^{i\varphi_m(\mathbf{u}_m)}F_m^*\left(\mathbf{u}_m\right)]. \quad (12)$$

При решении задачи минимизации функционала (3) с использованием градиентного метода расчет фазовых функций ДОЭ осуществляется итерационно. Опишем вычисления, происходящие на каждой итерации метода. Пусть $\varphi_1^k\left(\mathbf{u}_1\right),....,\varphi_n^k\left(\mathbf{u}_n\right)$ – фазовые функции ДОЭ, полученные на k-й итерации. Тогда для вычисления следующих приближений фазовых функций выполняются нижеприведенные шаги:

1) С использованием формул (1), описывающих прямое распространение поля, рассчитываются комплексные амплитуды полей $w_m(\mathbf{u}_m)e^{i\varphi_m(\mathbf{u}_m)}$ в плоскостях $z=f_m$, $m=1,\ldots,n$ и в выходной плоскости $z=f_{n+1}$.

2) При выбранном критерии оптимизации рассчитывается функция $F_{n+1}(\mathbf{u}_{n+1})$ в (6) и по формулам (7), (8), описывающим обратное распространение поля, рассчитываются функции $F_j(\mathbf{u}_j)$, $j=n, n-1,\ldots,1$.

3) По формуле (12) рассчитываются производные функционала $\delta\varepsilon/\delta\varphi_m$, $m=1,\ldots,n$.

4) Осуществляется расчет новых приближений фаз по формулам

$$\varphi_m^{k+1}\left(\mathbf{u}_m\right)=\varphi_m^k\left(\mathbf{u}_m\right)-t_k\frac{\delta\varepsilon}{\delta\varphi_m}\left(\mathbf{u}_m\right),\ m=1,...,n, \quad (13)$$

где t_k – шаг градиентного метода.

3. Случай нескольких падающих пучков

Представленный градиентный метод легко обобщается на задачу, в которой имеется $K>1$ различных входных распределений $w_{0,j}(\mathbf{u}_0)$, $j=1,\ldots,K$ (различных падающих пучков), и каскадный ДОЭ при каждом отдельном входном распределении должен сформировать в выходной плоскости «своё» заданное распределение интенсивности $I_j(\mathbf{u}_{n+1})$. В этом случае в качестве функционала ошибки можно использовать сумму функционалов

$$\varepsilon_{set}\left(\varphi_1,...,\varphi_n\right)=\sum_{j=1}^{K}\varepsilon^j\left(\varphi_1,...,\varphi_n\right), \quad (14)$$

где функционалы $\varepsilon^j(\varphi_1,\ldots,\varphi_n)$ представляют отличие распределений интенсивности $I_{n+1,j}(\mathbf{u}_{n+1})$, формируемых для входных распределений $w_{0,j}(\mathbf{u}_0)$, от требуемых распределений $I_j(\mathbf{u}_{n+1})$. Без ограничения общности можно считать, что данные функционалы имеют вид (2), но, возможно, с различными функциями $D[I_{n+1,j}(\mathbf{u}_{n+1}),I_j(\mathbf{u}_{n+1})]$. Поскольку производные от суммы функционалов (14) просто равны сумме производных слагаемых,

$$\frac{\delta\varepsilon_{set}\left(\varphi_1,...,\varphi_n\right)}{\delta\varphi_m}=\sum_{j=1}^{K}\frac{\delta\varepsilon^j\left(\varphi_1,...,\varphi_n\right)}{\delta\varphi_m},m=1,...,n, \quad (15)$$

то вычисление производных функционала (14) также сводится к формуле (12). При этом коррекция фазовых функций на каждой итерации осуществляется по формуле, аналогичной формуле (13).

Отметим, что рассмотренный градиентный метод также легко обобщается на задачу, когда требуемые распределения интенсивности $I_j(\mathbf{u}_{n+1})$ заданы в различных выходных плоскостях.

4. Расчет каскадных ДОЭ для решения задач классификации

В данном параграфе мы рассмотрим расчет каскадных ДОЭ для решения задачи классификации изображений. Опишем данную задачу. Пусть во входной плоскости $z=f=0$ располагается некоторый динамический транспарант, генерирующий амплитудные изображения объектов из L различных классов (например, изображения рукописных цифр от нуля до девяти). Транспарант освещается плоской волной с длиной волны λ. Формируемое таким образом световое поле распространяется через каскадный ДОЭ и попадает в выходную плоскость $z=f_{n+1}$. Будем считать, что в выходной плоскости заданы L пространственно разделенных областей G_j, соответствующих изображениям различных классов. Тогда задача расчета каскадного ДОЭ для решения задачи классификации состоит в расчете таких фазовых функций ДОЭ $\varphi_1(\mathbf{u}_1),\ldots,\varphi_n(\mathbf{u}_n)$, что при «входном сигнале», соответствующем изображению j-го класса, максимум формируемого распределения интенсивности достигается в соответствующей области G_j [17].

В задачах расчета каскадных ДОЭ для классификации изображений используется идеология нейронных сетей. В этом случае для расчета (обучения) каскадного ДОЭ используется т.н. обучающая выборка, содержащая набор входных распределений, соответствующих различным изображениям объектов различных классов. Вследствие большого размера обучающей выборки, обычно из неё случайным образом выбирается набор распределений (т.н. батч, от англ. batch) и для него рассчитываются производные функционала ошибки. Можно показать, что математические ожидания производных, рассчитанных по батчу, пропорциональны производным функционала, рассчитанным по всей выборке. В этой связи такой подход является методом стохастического градиента.

Процесс обучения каскадного ДОЭ на конкретном батче соответствует градиентному методу для

случая нескольких падающих пучков, рассмотренному в параграфе 3. Действительно, функционал ошибки при обучении по батчу можно считать заданным в виде (14), где K – размер батча, а функционалы $\varepsilon^j(\varphi_1,\ldots,\varphi_n)$ должны представлять ошибку распознавания распределений различных классов, входящих в батч. Отличие задачи классификации от задачи формирования различных распределений интенсивности для различных падающих пучков состоит только в виде функционалов $\varepsilon^j(\varphi_1,\ldots,\varphi_n)$. В следующих параграфах мы рассмотрим два функционала ошибки, используемых при решении задачи классификации, и покажем, что вычисление производных этих функционалов также сводится к общей формуле (12).

4.1. Квадратичный функционал ошибки

Пусть $w_{0,j}(\mathbf{u}_0)$ – входное распределение, соответствующее некоторому изображению j-го класса. Обозначим

$$\mathbf{u}_{n+1,k} = \arg\max_{\mathbf{u}_{n+1}\in G_k} I_{n+1}\left(\mathbf{u}_{n+1}\right), k = 1,\ldots,L, \tag{16}$$

координаты максимумов формируемого распределения интенсивности в областях G_k выходной плоскости при входном поле $w_{0,j}(\mathbf{u}_0)$. Для распознавания входного сигнала $w_{0,j}(\mathbf{u}_0)$ необходимо, чтобы в требуемой области G_j формировался максимум интенсивности с «большим» значением $I_{\max}$, а максимумы в остальных областях

$$I_{n+1}\left(\mathbf{u}_{n+1,k}\right) = \max_{\mathbf{u}_{n+1}\in G_k} I_{n+1}\left(\mathbf{u}_{n+1}\right)$$

были близки к нулю. В качестве значения $I_{\max}$ можно, например, использовать значение

$$I_{\max} = |(\lambda f_{n+1})^{-1}\iint w_{0,j}(\mathbf{u}_0)\, d^2\,\mathbf{u}_0|^2,$$

получающееся в фокусе при фокусировке входного изображения дифракционной линзой. Соответственно, в качестве функционала ошибки для распознавания входного распределения j-го класса можно использовать следующий квадратичный функционал

$$\begin{aligned}\varepsilon^j\left(\varphi_1,\ldots,\varphi_n\right) &= \sum_{k=1}^{L}\left(\max_{\mathbf{u}_{n+1}\in G_k} I_{n+1}(\mathbf{u}_{n+1}) - I_{\max}\delta_{k,j}\right)^2 = \\ &= \sum_{k=1}^{L}\left(I_{n+1}(\mathbf{u}_{n+1,k}) - I_{\max}\delta_{k,j}\right)^2,\end{aligned} \tag{17}$$

где $\delta_{k,j}$ – символ Кронекера.

Покажем, что вычисление производных функционала $\varepsilon^j(\varphi_1,\ldots,\varphi_n)$ аналогично вычислению производных «общего» функционала (2), рассмотренного в параграфе 2. Действительно, пусть $\Delta_m^j\varepsilon\left(\varphi_1,\ldots,\varphi_n\right)$ – приращение функционала (17), вызванное приращением $\Delta\varphi_m$ функции φ_m. Согласно (17), данное приращение имеет вид:

$$\begin{aligned}\Delta_m^j\varepsilon\left(\varphi_1,\ldots,\varphi_n\right) &= 2\sum_{k=1}^{L}\left(I_{n+1}(\mathbf{u}_{n+1,k}) - I_{\max}\delta_{k,j}\right)\Delta_m\left(w_{n+1}(\mathbf{u}_{n+1,k})w_{n+1}^*(\mathbf{u}_{n+1,k})\right) = \\ &= 4\,\mathrm{Re}\left[\sum_{k=1}^{L}\Delta_m w_{n+1}(\mathbf{u}_{n+1,k})\left(I_{n+1}(\mathbf{u}_{n+1,k}) - I_{\max}\delta_{k,j}\right)w_{n+1}^*(\mathbf{u}_{n+1,k})\right] = \\ &= 4\,\mathrm{Re}\left[\iint\Delta_m w_{n+1}(\mathbf{u}_{n+1})\left(I_{n+1}(\mathbf{u}_{n+1}) - I_{\max}\chi_{G_j}(\mathbf{u}_{n+1})\right)w_{n+1}^*(\mathbf{u}_{n+1})\sum_{k=1}^{L}\delta(\mathbf{u}_{n+1}-\mathbf{u}_{n+1,k})d^2\mathbf{u}_{n+1}\right] = \\ &= 4\,\mathrm{Re}\left\langle\Delta_m w_{n+1}(\mathbf{u}_{n+1}), F_{n+1}(\mathbf{u}_{n+1})\right\rangle,\end{aligned} \tag{18}$$

где $\delta(\mathbf{u}_{n+1})$ – дельта-функция, $\chi_{G_j}(\mathbf{u}_{n+1})$ – индикаторная функция множества G_j,

$$\begin{aligned}F_{n+1}\left(\mathbf{u}_{n+1}\right) &= \left(I_{n+1}(\mathbf{u}_{n+1}) - I_{\max}\chi_{G_j}(\mathbf{u}_{n+1})\right)\times \\ &\times w_{n+1}(\mathbf{u}_{n+1})\sum_{k=1}^{L}\delta(\mathbf{u}_{n+1}-\mathbf{u}_{n+1,k}).\end{aligned} \tag{19}$$

Аналогично (5), мы получили приращение функционала $\Delta_m^j\varepsilon\left(\varphi_1,\ldots,\varphi_n\right)$ в виде скалярного произведения. Соответственно, производные функционала $\varepsilon^j(\varphi_1,\ldots,\varphi_n)$ также задаются формулой (12), где функции $F_m(\mathbf{u}_m)$ рассчитываются через обратное распространение поля (19).

4.2. Функционал перекрестной энтропии

В задачах классификации в качестве критерия часто используют т.н. перекрестную энтропию в комбинации с функцией softmax для приведения максимумов интенсивности к диапазону [0,1] [17 – 19]. В этом случае для распознавания входного распределения j-го класса $w_{0,j}(\mathbf{u}_0)$ используется критерий

$$\varepsilon^j\left(\varphi_1,\ldots,\varphi_n\right) = -\ln\left[\frac{\exp(I_{n+1}(\mathbf{u}_{n+1,j}))}{\sum_{k=1}^{L}\exp(I_{n+1}(\mathbf{u}_{n+1,k}))}\right], \tag{20}$$

где $\mathbf{u}_{n+1,k}$ – координаты максимумов формируемого распределения интенсивности в областях G_k. Отметим, что выражение (20) будет близким к нулю, когда значение максимума интенсивности в требуемой области G_j будет существенно больше максимумов интенсивности в других областях.

Рассмотрим приращение функционала (20), вызванное приращением $\Delta\varphi_m$ функции φ_m. Проводя преобразования, аналогичные представленным выше, несложно получить приращение функционала в виде:

$$\Delta_m^j \varepsilon(\varphi_1,...,\varphi_n) = \\ = 4\,\mathrm{Re}\left[\iint \Delta_m w_{n+1}(\mathbf{u}_{n+1}) F_{n+1}^*(\mathbf{u}_{n+1}) d^2\mathbf{u}_{n+1}\right] = \\ = 4\,\mathrm{Re}\left\langle \Delta_m w_{n+1}(\mathbf{u}_{n+1}), F_{n+1}(\mathbf{u}_{n+1})\right\rangle, \quad (21)$$

где

$$F_{n+1}(\mathbf{u}_{n+1}) = \frac{w_{n+1}(\mathbf{u}_{n+1})}{2S(\mathbf{u}_{n+1})}[-S(\mathbf{u}_{n+1})\delta(\mathbf{u}_{n+1}-\mathbf{u}_{n+1,j}) + \\ + \exp(I_{n+1}(\mathbf{u}_{n+1}))\sum_{k=1}^{L}\delta(\mathbf{u}_{n+1}-\mathbf{u}_{n+1,k})], \quad (22)$$

где

$$S(\mathbf{u}_{n+1}) = \sum_{k=1}^{l}\exp(I_{n+1}(\mathbf{u}_{n+1}))\,.$$

Как и в предыдущем случае, мы записали приращение функционала $\Delta_m^j \varepsilon(\varphi_1,...,\varphi_n)$ в виде скалярного произведения. Соответственно, производные функционала (20) также задаются формулой (12), где функции $F_m(\mathbf{u}_m)$ рассчитываются через обратное распространение поля (22).

Таким образом, расчет фазовых функций каскадного ДОЭ в задаче классификации изображений состоит в следующем. Для текущего батча вычисляется градиент функционала (14), где расчет производных слагаемых производится по формулам (12), (19) или (12), (22), в зависимости от выбранного критерия. После вычисления производных функционала (14) осуществляется коррекция фазовых функций по формуле, аналогичной формуле (13). Затем рассматривается следующий батч и процесс повторяется.

5. Расчет каскадных ДОЭ для классификации рукописных цифр

В данном параграфе мы рассмотрим численный расчет ДОЭ для решения задачи классификации рукописных цифр. В качестве входных распределений будем использовать амплитудные изображения цифр из базы данных MNIST [25].

5.1. Расчет одиночных ДОЭ

Сначала рассмотрим решение задачи с помощью одного ДОЭ. Для расчета были выбраны следующие параметры. Входные изображения цифр задавались на сетке из 56×56 отсчетов с шагом $d = 18$ мкм. Фазовая функция ДОЭ задавалась на сетке 512×512 также с шагом $d = 18$ мкм. В этом случае размер стороны апертуры ДОЭ составляет 9,216 мм. Расстояния от входной плоскости до ДОЭ (d_1) и от ДОЭ до выходной плоскости (d_2) были выбраны одинаковыми и равными 300 мм. Отметим, что при расчетной длине волны $\lambda = 0532$ мкм угол дифракции на отсчете (пикселе) входного изображения равен $\varphi = \arcsin(\lambda/d) \approx 1{,}7°$. При этом картина дифракции от пикселя (по первому минимуму) на расстоянии $d_1 = 300$ мм примерно покрывает апертуру ДОЭ. Таким образом, выбранные параметры обеспечивают «связь» каждого пикселя входного изображения (нейрона входного слоя) со всеми отсчетами фазовой функции ДОЭ.

В соответствии с методом расчета в выходной плоскости $z = f_2 = 600$ мм были заданы 10 пространственно разделенных квадратных областей G_j с размером стороны 0,5 мм, в которых должны формироваться максимумы интенсивности для различных входных изображений различных цифр (см. рис. 2).

При расчете использовалась обучающая выборка, содержащая 60000 изображений цифр из базы данных MNIST. Расчет ДОЭ проводился с обучением по батчу (каждый батч содержал десять изображений, по одному для каждой цифры). В качестве функционалов ошибки использовались функционал квадратичной ошибки (КО) (14), (17) и критерий перекрестной энтропии (ПЭ) (14), (20). Расчет функций $w_m(\mathbf{u}_m)$, $F_m(\mathbf{u}_m)$, входящих в выражения для производных функционалов, был основан на численном расчете интегралов Френеля–Кирхгофа в (1), (7) с использованием процедуры быстрого преобразования Фурье. В качестве начального приближения для фазы ДОЭ использовалась нулевая фаза ($\varphi_1(\mathbf{u}_1) = 0$). Отметим, что использование случайной начальной фазы существенно не меняет процесс оптимизации.

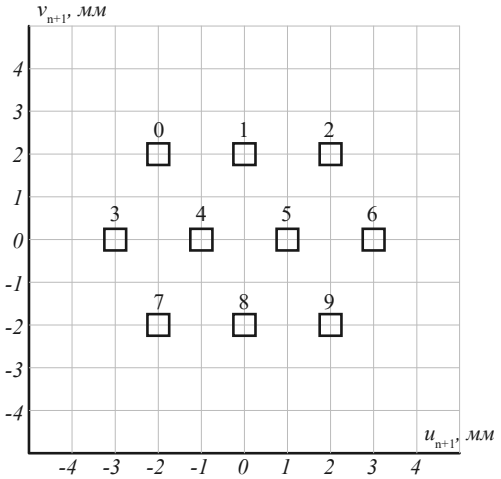


Рис. 2. Положение областей G_j в выходной плоскости, в которых должны формироваться максимумы интенсивности для входных изображений различных цифр

На рис. 3 для критериев КО и ПЭ представлены зависимости точности распознавания (процент правильно распознанных изображений), оцениваемые в процессе обучения (т.е. градиентной оптимизации) по последним обработанным 64 батчам (т.е. по последним 640 изображениям).

Из рис. 3 видно, что при использовании критерия КО точность распознавания быстро «выходит на плато» и примерно после 4000 итераций колеблется около значения примерно в 72 %. При использовании критерия ПЭ точность распознавания продолжает

возрастать и стабилизируется в районе 94 % примерно за 8000 итераций. Таким образом, рис. 3 показывает, что использование критерия ПЭ позволяет достичь лучшей точности распознавания. При этом время расчета ДОЭ для критерия ПЭ при выполнении 120000 итераций составило около 105 минут на стандартном персональном компьютере с процессором Intel® Core™ i9-10920X CPU @ 3.50GHz.

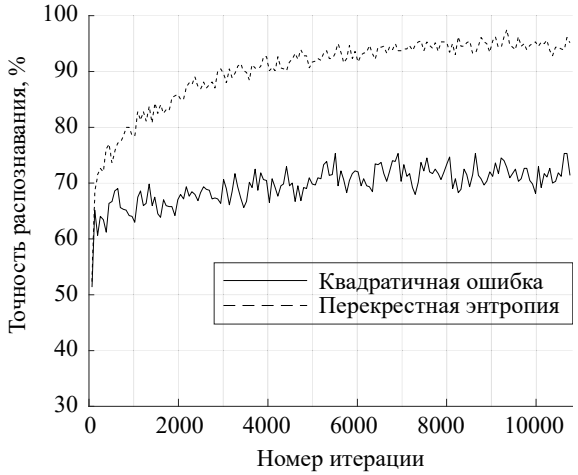
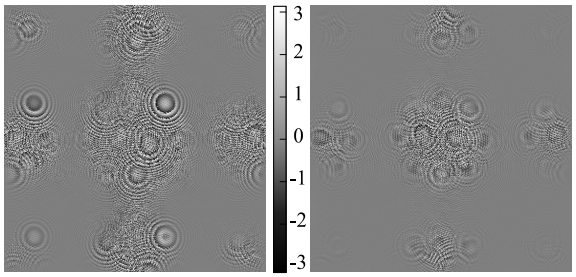


Рис. 3. Оценки точности распознавания для квадратичного критерия (непрерывная линия) и критерия перекрестной энтропии (штриховая линия) в зависимости от номера итерации (номера батча)

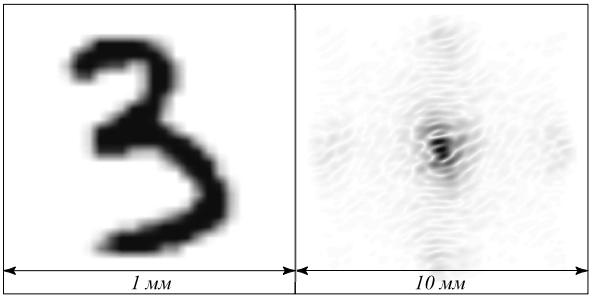
Фазовые функции ДОЭ, рассчитанные для обоих критериев, приведены на рис. 4. Можно видеть, что фазовые функции имеют области со значениями, близкими к нулю (к нулевой начальной фазе). Это связано с тем, что распределения амплитуд полей, формируемых в плоскости ДОЭ при входных изображениях в виде рукописных цифр, сосредоточены в центральной области и имеют несколько слабо выраженных максимумов в периферийных областях, между которыми поле близко к нулю.



а) *б)*

Рис. 4. Фазовые функции ДОЭ, рассчитанные с использованием критерия квадратичной ошибки (а) и критерия перекрестной энтропии (б)

В качестве примера на рис. 5 показаны входное изображение в виде цифры «3» и соответствующее распределение модуля комплексной амплитуды в плоскости ДОЭ. Можно увидеть явную зависимость между областями с малыми значениями фазы на рис. 4 и областями с малыми значениями поля на рис. 5*б*. Это связано с тем, что, согласно формуле (12), производная функционала ошибки оказывается близкой к нулю в областях с малым значением поля. Соответственно, в этих областях фазовая функция слабо изменяется и остается близкой к нулевой начальной фазе.



а) *б)*

Рис. 5. Входное распределение цифры «3» и распределение модуля комплексной амплитуды в плоскости ДОЭ

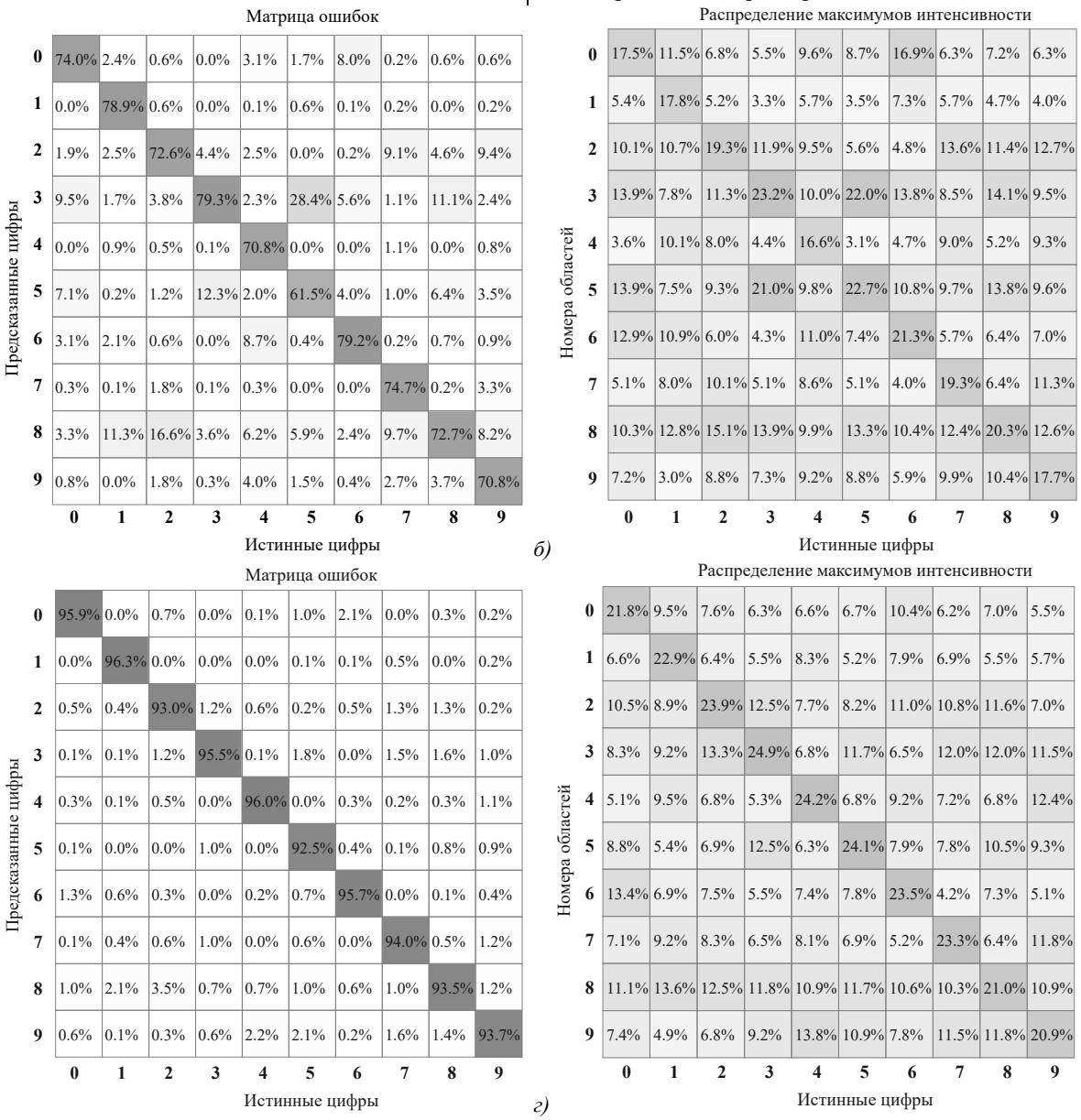
После процесса обучения выполнялось т.н. «слепое» тестирование работы рассчитанных ДОЭ (рис. 4) на тестовой выборке, содержащей 10000 изображений, которые не входили в обучающую выборку. В этом случае для каждого входного изображения из тестовой выборки рассчитывалось формируемое распределение интенсивности, вычислялись максимумы в областях G_j и затем по самому большому значению максимума определялась входная цифра. Результаты тестирования работы ДОЭ в виде матриц ошибок и «матриц интенсивности» представлены на рис. 6. Элемент (i,j) матрицы ошибок содержит процент распознаваний входных изображений цифры j как цифры i. Соответственно, диагональные элементы данных матриц содержат проценты правильных распознаваний. Аналогично, элемент (i,j) матрицы интенсивностей содержит усредненный максимум интенсивности (в процентах) в области G_j при входном изображении j-й цифры. При этом диагональные элементы матрицы интенсивности содержат относительную величину максимумов, выраженную в процентах, в требуемых областях для различных цифр.

Для ДОЭ, рассчитанного с использованием критерия КО (рис. 4*а*), точность распознавания (т.е. отношение правильно распознанных цифр к общему числу цифр в тестовой выборке) составляет 73,6 %. Для ДОЭ, рассчитанного с использованием критерия ПЭ (рис. 4*б*), точность распознавания (рис. 6*в*) значительно выше и составляет 94,6 %. Отметим, что достигнутая точность распознавания в 94,6 % является хорошей. Для сравнения отметим, что в работах [5, 21] теоретические точности распознавания, полученные для ДНН, состоящих из 5 и 10 ДОЭ, составили 92,3 % и 91,6 % соответственно. Указанные точности для каскадов ДОЭ значительно меньше достигнутой точности в 94,6 % для одного ДОЭ. При этом, как и в настоящей работе, расчет ДНН в [5, 21] проводился при длинах волн из видимого диапазона.

Важно отметить, что матрица интенсивностей для ДОЭ, рассчитанного по критерию ПЭ, имеет значительно лучший вид. Действительно, при практическом применении ДОЭ важной характеристикой является значение контраста, показывающее, насколько максимум интенсивности в требуемой области отличается от максимумов в других областях. Назовем контрастом для цифры i следующую величину:

$$\gamma_{\min,i} = \frac{I_{i,i} - \max\limits_{j \neq i} I_{j,i}}{I_{i,i} + \max\limits_{j \neq i} I_{j,i}}, \qquad (23)$$

где $I_{i,j}, i,j = 0,\ldots,9$ – элементы матрицы интенсивности. С учетом ошибок, неизбежно возникающих при экспериментальной реализации ДОЭ, желательно, чтобы в теоретических расчетах значения $\gamma_{\min,i}$ превосходили 0,1. Для матрицы интенсивностей на рис. 6*б* минимальный контраст достигается для цифры «5» и составляет всего $\gamma_{\min,5} \approx 0{,}06$. Для матрицы интенсивностей на рис. 6*г*, соответствующей ДОЭ, рассчитанному по критерию ПЭ, минимальный контраст достигается для цифры «9» и составляет $\gamma_{\min,9} \approx 0{,}25$. Таким образом, использование критерия ПЭ позволяет рассчитывать ДОЭ с существенно лучшими рабочими характеристиками.



а) *б)* *в)* *г)*

Рис. 6. Матрицы ошибок и матрицы интенсивности для ДОЭ, (а, б) рассчитанных по критерию квадратичной ошибки и (в, г) по критерию перекрестной энтропии

В качестве примера на рис. 7 показано распределение интенсивности, формируемое в выходной плоскости ДОЭ, рассчитанного по критерию ПЭ, при входном изображении в виде цифры «3» (рис. 5*а*). Можно видеть, что приведенное распределение интенсивности имеет выраженный максимум в требуе-

мой области G_3. Отметим, что данный максимум формируется вблизи границы области, поскольку использованный критерий оптимизации требует формирования максимума в соответствующей области без привязки к конкретной точке этой области.

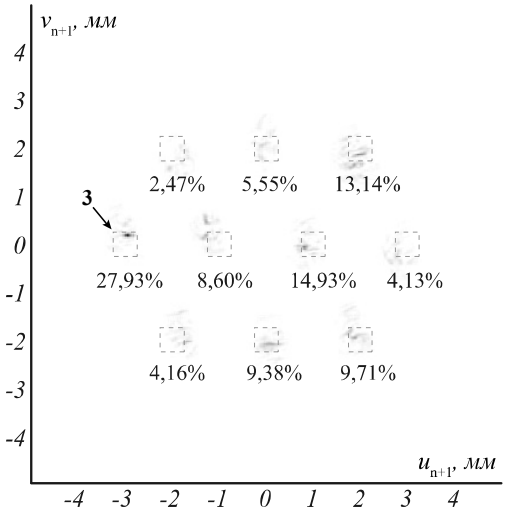
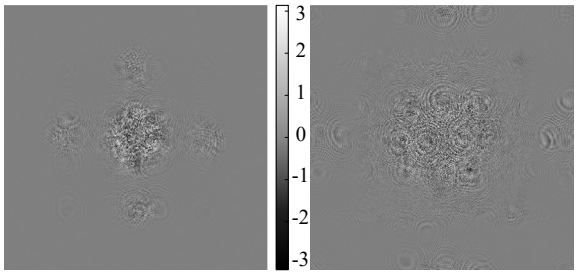


Рис. 7. Распределение интенсивности в выходной плоскости при входном изображении в виде цифры «3» на рис. 5а. Числовые значения указывают величины максимумов интенсивности в процентах в областях G_j

5.2. Расчет каскадного ДОЭ

Далее был рассчитан каскадный ДОЭ, состоящий из двух ДОЭ, расположенных в плоскостях $z=f_1=200$ мм и $z=f_2=400$ мм. Как и в предыдущих примерах, выходная плоскость расположена при $z=600$ мм. Все остальные параметры (дискретизация, длина волны, размеры апертур) совпадают с параметрами рассмотренных выше примеров.

Поскольку критерий ПЭ показал лучшие характеристики, то каскадный ДОЭ был рассчитан с использованием этого критерия. Расчет каскадного ДОЭ полностью аналогичен описанному выше расчету одиночных ДОЭ. На рис. 8 показаны фазовые функции каскадного ДОЭ, рассчитанные за 12000 итераций.



а) *б)*

Рис. 8. Фазовые функции каскадного ДОЭ в плоскостях z = 200 мм (а) z = 400 мм (б)

Результаты тестирования работы каскадного ДОЭ в виде матриц ошибок и матриц интенсивности представлены на рис. 9. Как и ранее, тестирование проводилось на выборке из 10000 изображений, которые не входили в обучающую выборку. Для рассчитанного каскадного ДОЭ точность распознавания составила 95,9 %, что всего на 1,3 % больше, чем для одиночного ДОЭ. В то же время матрица интенсивностей (рис. 9*б*) имеет значительно лучший вид, чем для одиночного ДОЭ. В частности, минимальный контраст, который также достигается для цифры «9», составляет 0,33 против 0,25 для одиночного ДОЭ. При этом средний контраст (среднее арифметическое контрастов для всех цифр) у каскадного ДОЭ составляет 0,37, что на 0,08 больше, чем для одиночного ДОЭ.

Отметим, что дальнейшее увеличение числа ДОЭ фактически не увеличивает точность распознавания, но несколько улучшает значения контраста. В частности, для трех ДОЭ точность распознавания составила 96 % (т.е. увеличилась всего на 0,1 %), а средний контраст составил 0,41 против 0,37 для случая двух ДОЭ.

Матрица ошибок

Предсказанные цифры \ Истинные цифры	0	1	2	3	4	5	6	7	8	9
0	98.2%	0.0%	0.3%	0.0%	0.3%	0.8%	1.0%	0.0%	0.2%	0.3%
1	0.0%	98.5%	0.2%	0.0%	0.0%	0.1%	0.3%	0.5%	0.0%	0.4%
2	0.1%	0.3%	93.6%	0.8%	0.9%	0.2%	0.4%	0.9%	0.2%	0.1%
3	0.2%	0.1%	1.0%	96.2%	0.1%	1.3%	0.0%	0.8%	0.9%	0.7%
4	0.0%	0.1%	1.0%	0.0%	95.0%	0.2%	0.6%	0.5%	0.9%	0.5%
5	0.1%	0.0%	0.1%	1.3%	0.0%	95.4%	1.6%	0.0%	2.0%	0.7%
6	0.4%	0.2%	0.3%	0.0%	0.5%	0.7%	95.5%	0.0%	0.2%	0.1%
7	0.3%	0.0%	0.9%	0.4%	0.1%	0.2%	0.1%	95.5%	0.5%	0.7%
8	0.7%	0.9%	2.5%	1.1%	0.7%	0.7%	0.4%	0.3%	94.5%	0.4%
9	0.0%	0.0%	0.2%	0.2%	2.3%	0.3%	0.0%	1.6%	0.6%	96.1%

а)

Распределение максимумов интенсивности

Номера областей \ Истинные цифры	0	1	2	3	4	5	6	7	8	9
0	29.5%	2.5%	6.0%	3.7%	4.0%	5.6%	8.6%	3.5%	5.1%	3.9%
1	2.3%	32.3%	5.7%	5.7%	4.0%	3.5%	4.7%	4.1%	4.3%	3.2%
2	9.9%	9.2%	28.6%	9.5%	6.4%	5.9%	11.9%	11.8%	10.5%	5.2%
3	7.1%	10.1%	11.7%	30.4%	5.2%	12.5%	5.5%	11.5%	12.0%	10.6%
4	6.4%	8.3%	5.5%	3.7%	33.5%	4.8%	10.8%	5.8%	6.3%	14.1%
5	10.0%	4.4%	6.1%	13.5%	4.7%	32.5%	9.7%	7.0%	11.0%	7.5%
6	8.2%	4.3%	7.4%	3.2%	6.3%	6.4%	28.1%	1.8%	5.2%	2.7%
7	5.6%	9.3%	8.8%	6.5%	8.6%	5.8%	3.6%	30.8%	6.1%	13.2%
8	11.2%	13.9%	13.4%	12.3%	10.0%	13.4%	9.5%	7.8%	26.8%	11.5%
9	9.7%	5.6%	6.7%	11.5%	17.3%	9.7%	7.5%	15.9%	12.6%	28.1%

б)

Рис. 9. Матрица ошибок (а) и матрица интенсивностей (б) для каскадного ДОЭ

Заключение

Рассмотрен градиентный метод расчета каскадного ДОЭ. С использованием свойства унитарности оператора распространения света через каскадный ДОЭ получены явные выражения для производных функционала ошибки общего вида по фазовым функциям каскадного ДОЭ. Рассмотрено применение градиентного метода в задаче фокусировки различных падающих пучков в области с различными распределениями интенсивности и в задаче классификации изображений. Представленные описания градиентного метода объединяют задачи синтеза каскадных ДОЭ для фокусировки лазерного излучения и для классификации изображений в рамках единого методологического подхода. При этом единственное отличие указанных задач состоит только в виде функционалов ошибки, производные которых, тем не менее, вычисляются на основе одной и той же общей формулы (12).

С использованием предложенного градиентного метода рассчитаны одиночные и каскадный ДОЭ для решения задачи классификации рукописных цифр. Представленные результаты численного моделирования демонстрируют хорошую точность распознавания, которая составила 94,6 % для одного ДОЭ и 95,9 % для каскада из двух ДОЭ. Для сравнения отметим, что теоретические точности распознавания, полученные в работах [3, 5, 21] для ДНН, состоящих от 5 до 10 ДОЭ, составляют от 91,6 % до 93,4 %.

По мнению авторов статьи, рассмотренный градиентный метод обобщается на задачу формирования заданных распределений интенсивности для нескольких падающих пучков с различными длинами волн. В этом случае в качестве «оптимизируемых параметров» функционала ошибки следует рассматривать не фазовые функции каскадного ДОЭ, а функции высоты микрорельефа. В остальном приведенные выкладки для расчета производных будут аналогичными. Кроме того, в силу установленной «общности» задач фокусировки и классификации, можно ожидать, что рассмотренный единый подход может быть также применен к решению различных задач классификации на различных длинах волн. Данные направления, посвященные расчету каскадных ДОЭ для работы с излучением различных длин волн, будут являться предметом дальнейших исследований.

Благодарности

Работа выполнена при поддержке Государственного задания ФНИЦ «Кристаллография и фотоника» РАН в части разработки градиентного метода расчета каскадных ДОЭ и Министерства науки и высшего образования Российской Федерации (государственное задание Самарскому университету, лаборатория «Фотоника для умного дома и умного города», проект FSSS-2021-0016) в части расчета ДОЭ для классификации изображений рукописных цифр.

References

[1] Zhang J, Pégard N, Zhong J, Adesnik H, Waller L. 3D computer-generated holography by non-convex optimization. Optica 2017; 4: 1306-1313.

[2] Wang H, Piestun R. Dynamic 2D implementation of 3D diffractive optics. Optica 2018; 5: 1220-1228.

[3] Lin X, Rivenson Y, Yardimci NT, Veli M, Luo Y, Jarrahiand M, Ozcan A. All-optical machine learning using diffractive deep neural networks. Science 2018; 361: 1004-1008.

[4] Schmidt S, Thiele S, Toulouse A, Bösel C, Tiess T, Herkommer A, Gross H, Giessen H. Tailored micro-optical freeform holograms for integrated complex beam shaping. Optica 2020; 7: 1279-1286.

[5] Zhou T, Fang L, Yan T, Wu J, Li Y, Fan J, Wu H, Lin X, Dai Q. In situ optical backpropagation training of diffractive optical neural networks. Photon Res 2020; 8: 940-953.

[6] Gerchberg R, Saxton W. A practical algorithm for the determination of phase from image and diffraction plane pictures. Optik 1972; 35: 237.

[7] Fienup JR. Phase retrieval algorithms: a comparison. Appl Opt 1982; 21: 2758-2769.

[8] Soifer VA, Kotlyar VV, Doskolovich LL. Iterative methods for diffractive optical elements computation. London: Taylor & Francis Ltd; 1997. ISBN: 0-7484-0634-4.

[9] Shechtman Y, Eldar YC, Cohen O, Chapman HN, Miao JW, Segev M. Phase retrieval with application to optical imaging. IEEE Signal Process Mag 2015; 32: 87-109.

[10] Latychevskaia T. Iterative phase retrieval in coherent diffractive imaging: practical issues. Appl Opt 2018; 57: 7187-7197.

[11] Ripoll O, Kettunen V, Herzig HP. Review of iterative Fourier transform algorithms for beam shaping applications. Opt Eng 2004; 43: 2549-2556.

[12] Doskolovich LL, Mingazov AA, Byzov EV, Skidanov RV, Ganchevskaya SV, Bykov DA, Bezus EA, Podlipnov VV, Porfirev AP, Kazanskiy NL. Hybrid design of diffractive optical elements for optical beam shaping. Opt Express 2021; 29(20): 31875-31890. DOI: 10.1364/OE.439641.

[13] Gülses AA, Jenkins BK. Cascaded diffractive optical elements for improved multiplane image reconstruction. Appl Opt 2013; 52: 3608-3616.

[14] Deng X, Chen RT. Design of cascaded diffractive phase elements for three-dimensional multiwavelength optical interconnects. Opt Lett 2000; 25: 1046-1048.

[15] Yan T, Wu J, Zhou T, Xie H, Xu F, Fan Jo, Fang L, Lin X, Dai Q. Fourier-space diffractive deep neural network. Phys Rev Lett 2019; 123: 023901.

[16] Zheng S, Xu S, Fan D. Orthogonality of diffractive deep neural network. Opt Lett 2022; 47: 1798-1801.

[17] Chang J, Sitzmann V, Dun X, Heidrich W, Wetzstein G. Hybrid optical-electronic convolutional neural networks with optimized diffractive optics for image classification. Sci Rep 2018; 8: 12324.

[18] Liu C, Ma Q, Luo ZJ, Hong QR, Xiao Q, Zhang HC, Miao L, Yu WM, Cheng Q, Li L, Cui TJ. A programmable diffractive deep neural network based on a digital-coding metasurface array. Nat Electron 2022; 5: 113-122.

[19] Mengu D, Luo Y, Rivenson Y, Ozcan A. Analysis of diffractive optical neural networks and their integration with electronic neural networks. IEEE J Sel Top Quantum Electron 2020; 26: 3700114.

[20] Sui X, Wu Q, Liu J, Chen Q, Gu G. A review of optical neural networks. IEEE Access 2020; 8: 70773-70783.

[21] Chen H, Feng J, Jiang M, Wang Y, Lin J, Tan J, Jin P. All-optical machine learning using diffractive deep neural networks. Engineering 2021; 361: 1483-1491.
[22] Kulce O, Mengu D, Rivenson Y, Ozcan A. All-optical synthesis of an arbitrary linear transformation using diffractive surfaces. Light Sci Appl 2021; 10: 196.
[23] Luo Y, Mengu D, Yardimci NT, Rivenson Y, Veli M, Jarrahiand M, Ozcan A. Design of task-specific optical systems using broadband diffractive neural networks. Light Sci Appl 2019; 8: 112.
[24] Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv Preprint. 2015. Source: <https://arxiv.org/abs/1412.6980>.
[25] Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE 1998; 86: 2278-2324.

Сведения об авторах

Сошников Даниил Вадимович, в 2022 году с отличием окончил бакалавриат Самарского национального исследовательского университета имени академика С.П. Королёва по специальности «Прикладная математика и информатика». E-mail: *soshnikov.d.v@mail.ru* .

Досколович Леонид Леонидович, в 1989 году с отличием окончил Куйбышевский авиационный институт (КуАИ, ныне – Самарский национальный исследовательский университет имени академика С.П. Королёва) по специальности «Прикладная математика». Доктор физико-математических наук (2001 год), профессор, главный научный сотрудник лаборатории дифракционной оптики Института систем обработки изображений РАН (ИСОИ РАН) — филиала ФНИЦ «Кристаллография и фотоника» РАН, профессор кафедры технической кибернетики Самарского университета и ведущий научный сотрудник научно-исследовательской лаборатории прорывных технологий дистанционного зондирования Земли Самарского университета. Специалист в области дифракционной оптики, лазерных информационных технологий, нанофотоники. E-mail: *leonid@ipsiras.ru* .

Бызов Егор Владимирович, 1988 года рождения. В 2014 году с отличием окончил обучение в магистратуре Самарского государственного аэрокосмического университета имени академика С.П. Королёва (СГАУ, ныне Самарский национальный исследовательский университет имени академика С.П. Королева) по направлению «Прикладные математика и физика». Кандидат физико-математических наук (2022 год), научный сотрудник лаборатории дифракционной оптики Института систем обработки изображений РАН – филиала ФНИЦ «Кристаллография и фотоника» РАН. Область научных интересов: методы расчетов формирующей неизображающей оптики для светодиодов. E-mail: *egor.byzov@gmail.com* .

ГРНТИ: 29.31.15
Поступила в редакцию 5 апреля 2023 г. Окончательный вариант – 31 мая 2023 г.

Gradient method for designing cascaded DOEs and its application in the problem of classifying handwritten digits

D.V. Soshnikov [1,2], L.L. Doskolovich [1,2], E.V. Byzov [1,2]
[1] IPSI RAS – Branch of the FSRC "Crystallography and Photonics" RAS, 443001, Samara, Russia, Molodogvardeyskaya 151;
[2] Samara National Research University, 443086, Samara, Russia, Moskovskoye Shosse 34

Abstract

We consider a gradient method for calculating cascaded diffractive optical elements (DOEs) consisting of several sequentially placed phase DOEs. Using the unitarity property of the operator describing the light propagation through the cascaded DOE, we obtained explicit expressions for the derivatives of the error functional with the respect to the phase functions of the cascaded DOE. We consider the application of the gradient method in the problem of focusing several different incident beams to several domains with different intensity distributions, and in the problem of image classification. The presented description of the gradient method treats the problems of designing cascaded DOEs for both focusing the laser radiation and performing image classification in the framework of a single general approach. It is shown that the difference of the problem of optical classification from the problem of generating required intensity distributions consists only in the form of error functionals, the calculation of the derivatives of which is reduced to the same general formula. Using the proposed gradient method, we designed single and cascaded DOEs for optical classification of handwritten digits. The obtained results may find application in the development of diffractive neural networks and optical systems for laser beam focusing.

Keywords: diffractive optical element, phase function, scalar diffraction theory, gradient method, image classification.

Citation: Soshnikov DV, Doskolovich LL, Byzov EV. Gradient method for designing cascaded DOEs and its application in the problem of classifying handwritten digits. Computer Optics 2023; 47(5): 691-701. DOI: 10.18287/2412-6179-CO-1314.

Acknowledgements: This work was performed within the State assignment of Federal Scientific Research Center "Crystallography and Photonics" of Russian Academy of Sciences in part of developing the gradient method for calculating cascaded DOEs, and was supported by the Ministry of Science and Higher Education of the Russian Federation (State assignment for research to Samara University (laboratory "Photonics for Smart Home and Smart City", project FSSS-2021-0016) in part of designing DOEs for classifying handwritten digits.

Authors' information

Daniil Vadimovich Soshnikov, graduated with honors (2022) from Samara National Research University with a major in Applied Mathematics and Computer Science. E-mail: *soshnikov.d.v@mail.ru* .

Leonid Leonidovich Doskolovich, graduated with honours (1989) from S.P. Korolyov Kuibyshev Aviation Institute (presently, Samara National Reseach University), majoring in Applied Mathematics. He received his Doctor in Physics & Maths (2001) degree from Samara State Aerospace University. Main reseacher of the Image Processing Systems Institute RAS – Branch of the FSRC "Crystallography and Photonics" RAS, professor at Technical Cybernetics department of National Research University, senior researcher at the Breakthrough Technologies for Earth's Remote Sensing laboratory at SSAU. His leading research interests include diffractive optics, laser information technologies, nanophotonics. E-mail: *leonid@ipsiras.ru* .

Egor Vladimirovich Byzov (b. 1988) graduated with honors (2014) from Samara State Aerospace University named after S.P. Korolyov (now – Samara National Research University named after academician S.P. Korolyov), majoring in Applied Mathematics and Physics. Candidate in Physics and Mathematics (2022). Currently he is a researcher at the Diffractive Optics Laboratory of the Image Processing Systems Institute (IPSI RAS – Branch of the FSRC "Crystallography and Photonics RAS"). Research interests: design methods of nonimaging optics for LEDs.
E-mail: *egor.byzov@gmail.com* .

Received April 5, 2023. The final version – May 31, 2023.