УДК 621.38

СИНТЕЗ ВОЛНОВЫХ ПОЛЕЙ В РАМКАХ ЭЛЕКТРОМАГНИТНОЙ ТЕОРИИ

© 1999 П.Г. Серафимович, С.И. Харитонов

Институт систем обработки изображений РАН, г. Самара

На протяжении последних лет большое внимание уделяется вопросам восстановления волнового поля по результатам измерения интенсивности в двух плоскостях. Подобного рода обратные задачи возникают в радиоастрономии, при расчетах дифракционных оптических элементов формирующих заданное распределение интенсивности, фокусаторах лазерного излучения. Последние часто используются для решения технологических проблем, связанных с лазерной обработкой материалов. Однако в подавляющем большинстве работ задачи восстановления полей решаются в рамках скалярной теории. В данной работе задача решается в рамках электромагнитной теории. Проведено численное сравнение результатов решения прямых и обратных задач в скалярном приближении и в рамках электромагнитной теории.

1. Алгоритм вычисления электромагнитного поля

В данном пункте будет рассмотрен метод расчета электромагнитных полей, основанный на численном решении уравнений Максвелла в импульсном представлении. Не ограничивая общности, рассмотрим распространение электромагнитного поля в среде, заполненной диэлектриком.

Систему уравнений Максвелла запишем в следующем виде:

$$\begin{aligned}
\partial_z E_x &= \frac{i}{k}\partial_x\left(\frac{1}{\varepsilon}\left(\partial_x H_y - \partial_y H_x\right)\right) + ikH_y \\
\partial_z E_y &= \frac{i}{k}\partial_y\left(\frac{1}{\varepsilon}\left(\partial_x H_y - \partial_y H_x\right)\right) - ikH_x \\
\partial_z H_x &= \frac{1}{ik}\partial_x\left(\partial_x E_y - \partial_y E_x\right) - ik\varepsilon E_y \\
\partial_z H_y &= \frac{1}{ik}\partial_y\left(\partial_x E_y - \partial_y E_x\right) + ik\varepsilon E_x
\end{aligned} \tag{1}$$

Данное представление удобно тем, что в нем используется только 2 компоненты векторов электрического и магнитного поля. Представим вектора электрического и магнитного поля и функцию диэлектрической проницаемости в виде разложения в интеграл Фурье по поперечным координатам[1-2]:

$$\begin{aligned}
\vec{E}(x,y,z) &= \int_{-\infty}^{\infty} \vec{E}(\alpha,\beta,z)\exp\left(ik\left(\alpha x + \beta y\right)\right) d\alpha d\beta \\
\vec{H}(x,y,z) &= \int_{-\infty}^{\infty} \vec{H}(\alpha,\beta,z)\exp\left(ik\left(\alpha x + \beta y\right)\right) d\alpha d\beta \\
\varepsilon(x,y,z) &= \int_{-\infty}^{\infty} \varepsilon(\alpha,\beta,z)\exp\left(ik\left(\alpha x + \beta y\right)\right) d\alpha d\beta \\
\varepsilon^{-1}(x,y,z) &= \int_{-\infty}^{\infty} \varepsilon^{-1}(\alpha,\beta,z)\exp\left(ik\left(\alpha x + \beta y\right)\right) d\alpha d\beta
\end{aligned} \tag{2}$$

Здесь символ ε^{-1} надо рассматривать, как единый символ т.е. $\varepsilon^{-1} \neq \frac{1}{\varepsilon}$.

Функции $\vec{E}(\alpha,\beta,z)$ и $\vec{H}(\alpha,\beta,z)$ в дальнейшем будем называть поперечными пространственно-частотными компонентами.

Подставляя эти разложения в уравнения Максвелла, записанные в форме (2), получаем уравнения Максвелла в импульсном (поперечном пространственно-частотном) представлении:

$$\partial_z E_x(\alpha,\beta,z) = -ik\alpha \int_{-\infty}^{\infty}\begin{pmatrix}\omega_x H_y(\omega_x,\omega_y,z) - \\ -\omega_y H_x(\omega_x,\omega_y,z)\end{pmatrix}\times$$
$$\times\varepsilon^{-1}\left(\alpha-\omega_x,\beta-\omega_y\right)d\omega_x d\omega_y + ikH_y(\omega_x,\omega_y,z)$$

$$\partial_z E_y(\alpha,\beta,z) = -ik\beta \int_{-\infty}^{\infty} \begin{pmatrix} \omega_x H_y(\omega_x,\omega_y,z) - \\ -\omega_y H_x(\omega_x,\omega_y,z) \end{pmatrix} \times$$
$$\times \varepsilon^{-1}\left(\alpha-\omega_x, \beta-\omega_y\right) d\omega_x d\omega_y + ikH_x(\omega_x,\omega_y,z)$$
$$\partial_z H_x\left(\alpha,\beta,z\right) = ik\alpha\left(\alpha E_y(\alpha,\beta,z) - \beta E_x(\alpha,\beta,z)\right) -$$
$$-ik\int_{-\infty}^{\infty} \varepsilon(\alpha-\omega_x,\beta-\omega_y) E_y(\omega_x,\omega_y,z) d\omega_x d\omega_y$$
$$\partial_z H_y\left(\alpha,\beta,z\right) = ik\beta\left(\alpha E_y(\alpha,\beta,z) - \beta E_x(\alpha,\beta,z)\right) + \tag{3}$$
$$+ik\int_{-\infty}^{\infty} \varepsilon(\alpha-\omega_x,\beta-\omega_y) E_x(\omega_x,\omega_y,z) d\omega_x d\omega_y$$

Приведенные формулы имеют громоздкий вид. Для дальнейшего анализа приведенных уравнений запишем их в операторно-матричной форме. Операторная форма записи уравнений позволит легко производить требуемые преобразования. В операторной записи система уравнений Максвелла имеет вид:

$$\partial_z \mathbf{W} = \mathbf{H}\mathbf{W}, \tag{4}$$

где $\mathbf{W}$ - матрица-столбец из 4-х компонент, $\mathbf{W} = \begin{pmatrix} E_x \\ E_y \\ H_x \\ H_y \end{pmatrix}$, $\mathbf{H}$-блочно-матричный дифференциальный оператор:

$$\mathbf{H} = \begin{pmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{B} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{5}$$

$$\mathbf{A} = \frac{i}{k} \begin{pmatrix} \partial_x & 0 \\ 0 & \partial_y \end{pmatrix} \begin{pmatrix} \varepsilon^{-1} & 0 \\ 0 & \varepsilon^{-1} \end{pmatrix} \begin{pmatrix} -\partial_y & \partial_x \\ -\partial_y & \partial_x \end{pmatrix} + ik \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

$$\mathbf{B} = \frac{1}{ik} \begin{pmatrix} \partial_x & 0 \\ 0 & \partial_y \end{pmatrix} \begin{pmatrix} -\partial_y & \partial_x \\ -\partial_y & \partial_x \end{pmatrix} + ik \begin{pmatrix} \varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Представление (2) в виде разложения по поперечным пространственно-частотным компонентам в матричной форме имеет вид:

$$\mathbf{W}(x,y.z) = \int_{-\infty}^{\infty} \mathbf{W}(\alpha,\beta,z) \exp\left(ik\left(\alpha x + \beta y\right)\right) d\alpha d\beta. \tag{6}$$

Система уравнений в пространственно-частотном представлении имеет вид:

$$\partial_z \mathbf{W}(\alpha,\beta,z) = \mathbf{H}(\alpha,\beta,z)\mathbf{W}(\alpha,\beta,z),$$

$$\mathbf{H}(\alpha,\beta,z) = \begin{pmatrix} \mathbf{A}(\alpha,\beta,z) & 0 \\ 0 & \mathbf{B}(\alpha,\beta,z) \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{7}$$

где двумерные матричные операторы действуют на двумерные матрицы-столбцы по формулам

$$\mathbf{A}(\alpha,\beta) \begin{pmatrix} \varphi \\ \psi \end{pmatrix} = ik \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \times$$
$$\times \int_{-\infty}^{\infty} \varepsilon^{-1}(\alpha-\omega_x,\beta-\omega_y,z) \begin{pmatrix} \omega_y & -\omega_x \\ \omega_y & -\omega_x \end{pmatrix} \times \tag{8}$$
$$\times \begin{pmatrix} \varphi(\omega_x,\omega_y,z) \\ \psi(\omega_x,\omega_y,z) \end{pmatrix} d\omega_x d\omega_y + ik \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \varphi \\ \psi \end{pmatrix}$$

$$\mathbf{B}(\alpha,\beta) \begin{pmatrix} \varphi \\ \psi \end{pmatrix} = ik \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \begin{pmatrix} \beta & -\alpha \\ \beta & -\alpha \end{pmatrix} \begin{pmatrix} \varphi \\ \psi \end{pmatrix} +$$
$$+ik\int_{-\infty}^{\infty} \varepsilon\left(\alpha-\omega_x,\beta-\omega_y\right) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \varphi(\omega_x,\omega_y) \\ \psi(\omega_x,\omega_y) \end{pmatrix} d\omega_x d\omega_y$$

Рассмотрим подробнее случай, когда электромагнитное поле распространяется в вакууме. В этом случае система интегро-дифференциальных уравнений превращается в систему обыкновенных дифференциальных уравнений:

$$\partial_z E_x(\alpha,\beta,z) = -ik\alpha\left(\alpha H_y(\alpha,\beta,z) - \beta\ \ H_x(\alpha,\beta,z)\right) +$$
$$+ikH_y(\alpha,\beta,z)$$
$$\partial_z E_y(\alpha,\beta,z) = -ik\beta\left(\alpha H_y(\alpha,\beta,z) - \beta H_x(\alpha,\beta,z)\right) -$$
$$-ikH_x(\alpha,\beta,z)$$
$$\partial_z H_x\left(\alpha,\beta,z\right) = ik\alpha\left(\alpha E_y(\alpha,\beta,z) - \beta E_x(\alpha,\beta,z)\right) -$$
$$-ikE_y(\alpha,\beta,z)$$
$$\partial_z H_y\left(\alpha,\beta,z\right) = ik\beta\left(\alpha E_y(\alpha,\beta,z) - \beta E_x(\alpha,\beta,z)\right) +$$
$$+ikE_x(\alpha,\beta,z) \tag{9}$$

Общее решение данной системы уравнений имеет вид:

$$\mathbf{W}(\alpha,\beta,z)=\begin{pmatrix} E^{+}(\alpha,\beta)\mathbf{W}^{+e}(\alpha,\beta)+ \\ +H^{+}(\alpha,\beta)\mathbf{W}^{+\mathbf{h}}(\alpha,\beta)\end{pmatrix}\times$$
$$\times\exp\left(ik\sqrt{1-\alpha^{2}-\beta^{2}}z\right)+$$
$$+\begin{pmatrix} E^{-}(\alpha,\beta)\mathbf{W}^{-\mathbf{e}}(\alpha,\beta)+ \\ +H^{-}(\alpha,\beta)\mathbf{W}^{-\mathbf{h}}(\alpha,\beta)\end{pmatrix}\times$$
$$\times\exp\left(-ik\sqrt{1-\alpha^{2}-\beta^{2}}z\right)$$

$$\|W\|\mathbf{W}^{\pm\mathbf{e}}(\alpha,\beta)=\begin{pmatrix}\mp\sqrt{1-\alpha^{2}-\beta^{2}}\alpha \\ \mp\sqrt{1-\alpha^{2}-\beta^{2}}\beta \\ \beta \\ -\alpha\end{pmatrix},$$

$$\|W\|\mathbf{W}^{\pm\mathbf{h}}(\alpha,\beta)=\begin{pmatrix}-\beta \\ \alpha \\ \mp\sqrt{1-\alpha^{2}-\beta^{2}}\alpha \\ \mp\sqrt{1-\alpha^{2}-\beta^{2}}\beta\end{pmatrix}, \qquad (10)$$

где $\|W\|=\dfrac{1}{\sqrt{\left(2-\alpha^{2}-\beta^{2}\right)\left(\alpha^{2}+\beta^{2}\right)}}$.

Верхний знак соответствует волнам, распространяющимся в положительном направлении, а нижний знак в отрицательном. Для полноты изложения отметим, что переход к случаю однородной диэлектрической среды с показателем преломления ε осуществляется с помощью следующего преобразования

$$\alpha\to\frac{\alpha}{\sqrt{\varepsilon}},\quad \beta\to\frac{\beta}{\sqrt{\varepsilon}},\quad z\to\sqrt{\varepsilon}z,$$

$$H_x\to\sqrt{\varepsilon}\,H_x,\ H_y\to\sqrt{\varepsilon}\,H_y. \qquad (11)$$

Перепишем полученное выражение в блочно-матричной форме

$$W(\alpha,\beta,z)=\mathbf{WSP}, \qquad (10)$$

$$\mathbf{P}=\begin{pmatrix}E^{+}(\alpha,\beta) \\ H^{+}(\alpha,\beta) \\ E^{-}(\alpha,\beta) \\ H^{-}(\alpha,\beta)\end{pmatrix},$$

$$\mathbf{S}=\begin{pmatrix}\exp\left(ik\sqrt{1-\alpha^{2}-\beta^{2}}z\right)E_{2} & 0 \\ 0 & \exp\left(-ik\sqrt{1-\alpha^{2}-\beta^{2}}z\right)E_{2}\end{pmatrix},$$

$$\mathbf{W}=\left(W^{+e}(\alpha,\beta)\quad W^{+h}(\alpha,\beta)\quad W^{-e}(\alpha,\beta)\quad W^{-h}(\alpha,\beta)\right),$$

$$\mathbf{E}=\begin{pmatrix}1 & 0 \\ 0 & 1\end{pmatrix}.$$

Связь между пространственно-частотными компонентами в различных плоскостях имеет вид:

$$W(\alpha,\beta,z)=\mathbf{WSM}^{-1}\overline{\mathbf{W}}W(\alpha,\beta,0), \qquad (12)$$

где $\overline{\mathbf{W}}$-матрица эрмитово-сопряженная к $\mathbf{W}$,

$$\mathbf{M}=\begin{pmatrix}E_{2} & W^{\pm}E_{2} \\ W^{\pm}E_{2} & E_{2}\end{pmatrix},$$

$$W^{\pm}=\overline{W}^{+e}W^{-e}=\frac{\left(\alpha^{2}+\beta^{2}\right)^{2}}{\|W\|^{2}}$$

Используя вышеприведенные выражения, подставляя в (6), получим выражение оператора распространения электромагнитного поля в свободном пространстве (пропагатора):

$$W(x,y,z)=TW(x,y,0)=\frac{k^{2}}{4\pi^{2}}\times$$
$$\times\int_{-\infty}^{\infty}\mathbf{WSM}^{1}\overline{\mathbf{W}}\begin{pmatrix}\int_{-\infty}^{\infty}W(x©,y©,0)\times \\ \times\exp\left(ik\begin{pmatrix}\alpha(x-x©)+ \\ +\beta(y-y©)\end{pmatrix}\right)dx©dy©\end{pmatrix}d\alpha d\beta$$

(13)

Преимущество данного алгоритма состоит в том, что в его основе лежит вычисление прямого и обратного преобразования Фурье. Это позволяет построить быстрые преобразования, основанные на алгоритмах БПФ, что в свою очередь позволяет значительно сократить время вычислений по сравнению с решением уравнений Максвелла разностными методами.

2. Определение псевдоскалярного произведения и унитарность оператора распространения

Введем в пространстве четырехкомпонентных векторов функций W понятие скалярного произведения. Введем скалярное произведение таким образом, чтобы оно не зависело от координаты z, и в тоже время произведение вектора самого на себя было пропорционально потоку вектора Умова-Пойтинга. Для этого запишем уравнения Максвелла в обычной форме для поля $E_1\, H_1$ и для комплексно сопряженного поля $E^*{}_2\, H^*{}_2$:

$$
\begin{aligned}
&rotE_1 = ikH_1\\
&rotH_1 = -ikE_1\\
&rot\,E^*{}_2 = -ikH^*{}_2\,,\\
&rot\,H^*{}_2 = ikE^*{}_2
\end{aligned}
\qquad (14)
$$

вычитаем второе уравнение из первого, и, используя известную формулу векторного анализа

$$H^*{}_2 rotE_1 - E_1 rotH^*{}_2 = ikH_1H^*{}_2 - ikE_1E^*{}_2\,,$$

$$E^*{}_2 rotH_1 - H_1 rot\,E^*{}_2 = -ikE_1E^*{}_2 + ikH_1H^*{}_2\,,$$

$$div[a,b] = brota - arotb\,, \qquad (15)$$

получаем следующее выражение:

$$div[E_1, H^*{}_2] - div[H_1, E^*{}_2] = \qquad (16)$$

или $div\big([E_1, H^*{}_2] + [E^*{}_2, H_1]\big) = 0$.

Далее, используя теорему Остроградского-Гаусса и учитывая условия излучения (с целью зануления интеграла по боковой поверхности), получаем:

$$
\iint \begin{pmatrix} [E_1(x,y,z_1), H^*{}_2(x,y,z_1)] + \\ + [E^*{}_2(x,y,z_1), H_1(x,y,z_1)] \end{pmatrix} dxdy =
$$

$$
= \iint \begin{pmatrix} [E_1(x,y,z_2), H^*{}_2(x,y,z_2)] + \\ + [E^*{}_2(x,y,z_2), H_1(x,y,z_2)] \end{pmatrix} dxdy = . \qquad (17)
$$

Запишем последнее выражение в следующем виде:

$$\iint A^T((x,y,z_1))\Omega B^*(x,y,z_1)dxdy =$$

$$\iint A^T((x,y,z_2))\Omega B^*(x,y,z_2)dxdy \quad , \qquad (18)$$

где

$$
\Omega = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}
$$

или

$$\iint < A(x,y,z_2), B(x,y,z_2) > dxdy =$$

$$\iint < A(x,y,z_1), B(x,y,z_1) > dxdy \quad , \qquad (19)$$

где $< A,B >= \left(A^T\right)^* \Omega B$ - псевдоскалярное произведение в пространстве четырехкомпонентных матриц столбцов.

Формула (19) означает, что оператор распространения (13) сохраняет скалярное произведение, т.е. является в пространстве с данным скалярным произведением унитарным оператором. Это свойство сохранения скалярного произведения можно использовать для решения обратных задач дифракции. Это в значительной мере упрощает вычисление градиента функционала невязки по аналогии с тем, как это наблюдается в случае скалярного приближения [3-4]. Кроме того, наличие сохраняющейся величины можно использовать для контроля правильности решения прямой задачи дифракции.

3. Пример численного моделирования

В качестве примера приведем вычисление поля от геометрооптического фокусатора гауссова пучка в тонкое кольцо и в квадрат. В этом случае поле в плоскости непосредственно в плоскости за оптическим элементом имеет вид:

$$
W(x,y,0) = \frac{1}{2}\sqrt{I_0(x,y)}\exp\big(i\varphi(x,y)\big)\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \qquad (20)
$$

где $I_0(x,y)$ - интенсивность падающего пучка, $\varphi(x,y)$ - фазовая функция геометрооптического фокусатора. При вычислении поля были использованы следующие пара-

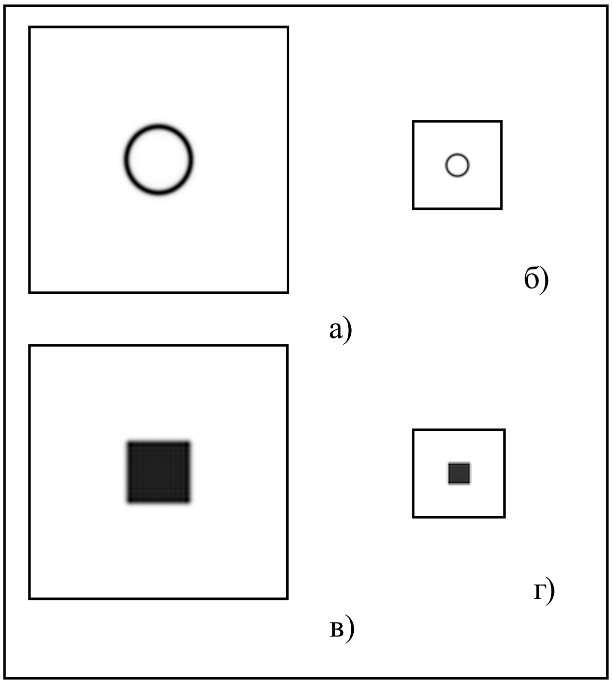


Рис. 1. Распределение интенсивности от геометрооптических фокусаторов

метры: размер ДОЭ по горизонтали и вертикали, соответственно: $D_x = 256\lambda$, $D_y = 256\lambda$, параметр гауссова пучка $\sigma = 64\lambda$, фокусное расстояние $f = 512\lambda$.

Кроме того, для Рис.1: (а) шаг дискретизации $\Delta = \lambda$, радиус кольца $r_0 = 32\lambda$; (б) шаг дискретизации $\Delta = 4\lambda$, радиус кольца $r_0 = 32\lambda$; (в) шаг дискретизации $\Delta = \lambda$, размер квадрата $a \times b = 32\lambda \times 32\lambda$; (г) шаг дискретизации $\Delta = 4\lambda$, размер квадрата $a \times b = 32\lambda \times 32\lambda$.

Полученные результаты говорят о том, что предлагаемый метод расчета дает результат, согласующийся со скалярной теорией.

Проведем сравнение численных результатов расчета поля от геометрооптического фокусатора гауссова пучка в квадрат, полученных в приближении Френеля и в рамках электромагнитной теории, при следующих параметрах:

шаг дискретизации на элементе $\Delta = \lambda$, размеры апертуры элемента $D_x = 64\lambda$, $D_y = 64\lambda$, фокусное расстояние $f = 64\lambda$, размер квадрата в плоскости фокусировки, $a \times b = 16\lambda \times 16\lambda$. Кроме того, для рис.2: (а) $\sigma = 16\lambda$ - параметр гауссова пучка, дифракционная эффективность eff = 93%, среднеквадратичное отклонение рассчитанного распределения интенсивности от заданного rms = 20% (электромагнитная теория); (б) $\sigma = 16\lambda$, eff = 94%, rms = 17% (скалярная теория); (в) $\sigma = 32\lambda$, eff = 96%, rms = 49% (электромагнитная теория); (г) $\sigma = 32\lambda$, eff = 91%, rms = 26% (скалярная теория).

Анализ полученных результатов показывает, что использование строгой электромагнитной теории для решения прямой задачи, т.е. вычисления поля, не приводит к качественным отличиям по сравнению с результатами, полученными в скалярном приближении. Отличие результатов, полученных в рамках скалярной и векторной теории, зависят от параметров задачи. Наиболее сильные отличия наблюдаются, когда размер освещающего пучка больше апертуры ДОЭ.

4. Градиентный метод синтеза волновых полей

Рассмотрим метод восстановления волновых полей по результатам измерения интенсивности электромагнитного поля в двух различных плоскостях. В качестве интенсивности в данном случае выступает величина, равная проекции вектора Умова-Пойнтинга на ось z декартовой системы координат.

Пусть $I_0(x, y)$ - распределение интенсивности в плоскости z = 0, $I(x, y)$ - распре-

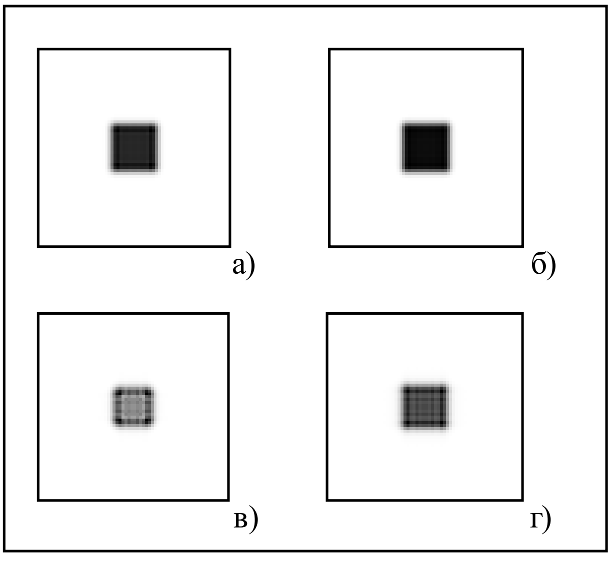


Рис. 2. Распределение интенсивности от фокусатора в квадрат при различных параметрах

деление интенсивности в плоскости z = f. Требуется найти волновое поле $W(x,y,z)$, удовлетворяющее следующим условиям:

$$\begin{aligned} &W(x,y,f) = TW(x,y,0) \\ &I_0(x,y) = < W(x,y,0), W(x,y,0) >, \\ &I\ (x,y) = < W(x,y,f), W(x,y,f) >, \end{aligned} \qquad (21)$$

где T - оператор распространения.

Точное решение задачи в общем случае не существует, поэтому заменим точную постановку задачи приближенной:

$$\min \varepsilon\left[W(x,y,f)\right], \qquad (22)$$

$$\varepsilon = \int \left(\sqrt{I(x,y)} - \sqrt{< W(x,y,f), W(x,y,f) >} \right)^2 dxdy,$$

$$W(x,y,f) = TW(x,y,0),$$

$$I_0(x,y) = < W(x,y,0), W(x,y,0) >.$$

Для простоты будем искать минимум в классе функций, имеющих следующий вид:

$$W(x,y,0) = \frac{1}{2}\sqrt{I_0(x,y)}\exp\left(i\varphi(x,y)\right)E(x,y,0), \qquad (23)$$

$$E(x,y,0) = \begin{pmatrix} \cos\alpha(x,y) \\ -\sin\alpha(x,y)\exp\left(i\delta(x,y)\right) \\ -\sin\alpha(x,y)\exp\left(i\delta(x,y)\right) \\ \cos\alpha(x,y) \end{pmatrix}.$$

Отметим, что при таком выборе условие (21) выполняется автоматически. Этот вид представления поля выбран потому, что при данной форме записи каждый из параметров в приближении геометрической оптики имеет простой физический смысл. Функция $\varphi(x,y)$ определяет фазу (или эйконал) волны в плоскости z = 0, $\alpha(x,y), \delta(x,y)$ определяют поляризацию волны.

Найдем градиент функционала

$$\delta\varepsilon = 0.5\left(\begin{aligned} &\left\langle \delta W(\vec{x},f), \sqrt{\frac{I(\vec{x})}{\langle W(\vec{x},f), W(\vec{x},f)\rangle}} W(\vec{x},f) \right\rangle + \\ &+ \left\langle \sqrt{\frac{I(\vec{x})}{\langle W(\vec{x},f), W(\vec{x},f)\rangle}} W(\vec{x},f), \delta W(\vec{x},f) \right\rangle \end{aligned} \right) =,$$

$$= 0.5\left(\langle \delta W(\vec{x},0), \psi(\vec{x},0)\rangle + \langle \psi(\vec{x},0), \delta W(\vec{x},0)\rangle \right) \qquad (24)$$

где

$$\psi(x,0) = T^{-1}\left(\sqrt{\frac{I(\vec{x})}{\langle W(\vec{x},f), W(\vec{x},f)\rangle}} W(\vec{x},f) \right),$$

$$\delta W(\vec{x},f) = T(\delta W(\vec{x},0)).$$

При выводе формулы (24) было использовано свойство сохранения псевдоскалярного произведения. Переходя к дискретному представлению и используя вышеприведенные выражения, получим выражение для градиента функционала в каждой точке (u, v) по аналогии с [1-2].

$$\nabla\varepsilon(\vec{x},0) = \langle L(\vec{x},0), \psi(\vec{x},0)\rangle + \langle \psi(\vec{x},0), L(\vec{x},0)\rangle, \qquad (25)$$

где $L(\vec{x},0) = \dfrac{i}{2}\sqrt{I_0(\vec{x},0)}\exp(i\varphi(\vec{x},0))E(\vec{x},0)$.

Далее, используя выражение для градиента, можно построить итерационную процедуру минимизации функционала $\varphi_{n+1}(\vec{x},0) = \varphi_n(\vec{x},0) - t\nabla\varepsilon$, t-шаг градиентного алгоритма.

Полученное решение сильно зависит от начального приближения. Это связано с тем, что функционал обычно не имеет глобального минимума, имеет много локальных минимумов. Для решения задачи восстановления полей это - существенный недостаток, и для решения задачи в большинстве случаев необходимо использование дополнительной информации о восстанавливаемом поле, например, степени гладкости или форме спектра поля.

Данный метод можно использовать как для восстановления волновых полей, так и для их синтеза. Последнее используется при расчете дифракционных оптических элементов. В этом случае не возникает проблем с множеством решения задачи, так как необходимо получить любое решение, минимизирующее исходный функционал.

5. Пример численного решения задачи синтеза

В качестве примера приведем расчет поля на выходе дифракционного оптического элемента (ДОЭ), фокусирующего в квадрат, расположенный на расстоянии f от плоскости ДОЭ (рис3). Шаг дискретизации

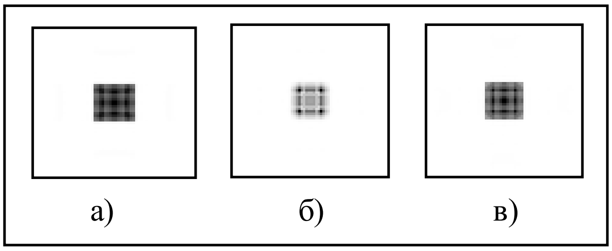


Рис. 3. Распределение интенсивности от фокусатора в квадрат, рассчитанного итерационным методом в различных приближениях

на элементе $\Delta = \lambda$, размеры апертуры элемента $D_x = 64\lambda$, $D_y = 64\lambda$, $f = 64\lambda$ - фокусное расстояние, размер квадрата в плоскости фокусировки, $a \times b = 16\lambda \times 16\lambda$, $\sigma = 32\lambda$ - параметр гауссова пучка.

На рисунке 3а изображено распределение интенсивности от фокусатора гауссова пучка в квадрат, рассчитанное в приближении Френеля. Расчет поля также приведен в приближении Френеля. Дифракционная эффективность eff=94%, среднеквадратичное отклонение rms=16%.

На рис 3б изображено распределение интенсивности от фокусатора гауссова пучка в квадрат, но расчет поля проводился в рамках электромагнитной теории, eff=96%, rms=64%.

На рис 3в изображено распределение интенсивности от фокусатора гауссова пучка в квадрат, рассчитанное в рамках электромагнитной теории, eff=97%, rms=22%.

Полученные результаты свидетельствуют о том, что использование строгой электромагнитной теории в рассмотренных задачах качественно не меняет результат решения обратной задачи. Количественные отклонения в большей степени зависят от параметров задачи. Количественный анализ полученных результатов показывает, что результаты, полученные в скалярном приближении и в рамках электромагнитной теории, отличаются в случае, если размер освещающего пучка превышает размер апертуры ДОЭ. В остальных случаях они совпадают с точностью до нескольких процентов.

СПИСОК ЛИТЕРАТУРЫ

1. Electromagnetic Theory of Gratings: Topics in current physics, v.22, Ed. by *R.Petit, N.Y.: Springer-Verlag*, 1980.
2. *В.А. Зверев*. Радиооптика. М. Советское Радио, 1975.
3. *Воронцов М.А., Корябин А.В., Шмальгаузен В.И.* Управляемые оптические системы. М.:Наука, 1988.- 272с.
4. *Воронцов М.А., Шмальгаузен В.И.* Принципы адаптивной оптики.- М.: Наука, 1985.-335с.

SYNTHESIS OF WAVE FIELDS IN FRAMEWORK OF ELECTROMAGNETIC THEORY

© 1999 P.G. Serafimovich, S.I. Kharitonov

Image Processing System Institute of Russian Academy of Sciences, Samara

In the last decade the issue of wave field determination on the base of intensity measurements in two planes is at the centre of attention. Such inverse problems arise in radio-astronomy; during calculation of diffraction optical elements, which form a specified intensity distribution; and for laser radiation focusing device. The latter is often used for dealing with technological problems, which are connected with laser treatment of materials. However, the problems of wave fields determination are solved in the framework of scalar theory mostly. In this paper the problem is solved in the framework of electromagnetic theory. The results of direct and inverse problems solving for scalar approximation and within the framework of electromagnetic theory were numerically compared in the paper.

УДК 621.395.4, 621.372.542

БЫСТРЫЕ АЛГОРИТМЫ ДИСКРЕТНЫХ КОСИНУСНЫХ ПРЕОБРАЗОВАНИЙ КОРОТКИХ ДЛИН С МИНИМАЛЬНОЙ ВЫЧИСЛИТЕЛЬНОЙ СЛОЖНОСТЬЮ

© 1999 В. М. Чернов, М. А. Чичева

Институт систем обработки изображений РАН, г. Самара

В работе рассматриваются быстрые алгоритмы дискретного косинусного преобразования (ДКП) коротких длин с минимальной вычислительной сложностью. Снижение вычислительной сложности достигается за счет применения нового подхода к синтезу алгоритмов ДКП коротких длин, связанного с интерпретацией вычисления ДКП как операций в ассоциированных алгебраических структурах. Исследуется применение разработанных алгоритмов в методе блочного кодирования изображений.

Введение

Дискретное косинусное преобразование (ДКП)[1]:

$$C_x(m)=\lambda_m\sum_{n=0}^{N-1}x(n)\cos\left(\frac{\pi m(2n+1)}{2N}\right),$$

$$(m = 0, ..., N\text{-}1), \quad (1)$$

где

$$C_x(m)=\lambda_m\sum_{n=0}^{N-1}x(n)\cos\left(\frac{\pi m(2n+1)}{2N}\right)$$

- нормировочные коэффициенты [1], широко используется в обработке изображений. В частности, ДКП длины N=8 является основой целого ряда современных стандартов кодирования (JPEG, MPEG, ITU-T [2, 3]). Однако метод блочного кодирования с преобразованием (на котором базируются перечисленные стандарты) не в полной мере учитывает особенности конкретного изображения. Попытки построения *адаптивных* алгоритмов кодирования на основе ДКП [4, 5, 6, 7] связаны, обычно, с резким увеличением вычислительной сложности ДКП при возрастании длины преобразования. Это связано с тем, что быстрые алгоритмы (БА) ДКП при N>8 изучены в значительно меньшей степени, чем ДКП длины 8, которому посвящено множество работ [1, 2, 8, 9, 10].

Большинство из известных БА ДКП, синтезированных для преобразований произвольных длин обладают относительно низкой *асимптотической* арифметической сложностью, но не учитывают реализационной специфики «коротких» длин: относительно небольшого числа различных значений базисных функций, высоких требований к *структурной* простоте, которая при небольшой длине преобразования является определяющей характеристикой быстродействия и т.д.

Например, в [1] предлагается метод сведения ДКП к дискретному преобразованию Фурье (ДПФ) вещественной последовательности двойной длины:

$$\tilde{x}(m)=\frac{1}{2}\omega^{m/2}\sum_{k=0}^{2N-1}y(k)\omega^{mk},\quad \omega=\exp\left\{\frac{2\pi i}{2N}\right\} \quad (2)$$

где

$$y(k)=\begin{cases}x(k) & \text{при}\quad 0\le k\le N-1\\ x(2N-k-1) & \text{при}\quad N\le k\le 2N-1\end{cases}$$

Вычисление преобразования (2) для N=8 и при использовании одного из лучших БА ДПФ (сплит-радикс алгоритма вещественного ДПФ [11]) длины 16 требует согласно оценкам работы [12] 34 операции вещественного умножения и 56 операций вещественного сложения, в то время как специфический алгоритм работы [8] требует всего 12 умножений и 29 сложений. Алгоритмы косинусного преобразования «нестандартных» длин $N\neq 2^k$ [13, 14] также не ориентированы на работу с «короткими» ДКП.

Целью настоящей работы является разработка алгоритмической поддержки новых систем адаптивного блочного кодирования в виде высокоскоростных БА ДКП для обработки двумерных массивов малых объемов (использование блоков $N\times N$ при $8\le N<16$). Снижение вычислительной сложности БА ДКП (а, следовательно, и времени сжатия/восстановления изображения) достигается с помощью нового подхода к синтезу таких алгоритмов. Принципы такого подхода, связанного с интерпретацией вычисления ДКП как операций в ассоциированных алгебраических структурах (конечномерных алгебрах) были описаны в [15] (см. также [16, 17]). В настоящей статье на основе методики работы [15] авторы синтезируют алгоритмы ДКП с неулучшаемыми оценками вычислительной сложности.

1. Алгебраические принципы синтеза БА ДКП коротких длин

Рассматриваемый метод синтеза быстрых алгоритмов дискретного косинусного преобразования базируется на следующих алгебраических идеях.

1. Матрица ДКП имеет *блочную структуру*. Результат умножения такой матрицы на входной вектор сводится к умножению векторов из подпространств сигнального пространства на матрицы меньших размеров со специфическими свойствами «симметрии».

2. Умножение этих подматриц на векторы соответствующих подпространств эквивалентно умножениям элементов некоторых *конечномерных алгебр*.

3. В большинстве (но не во всех) рассматриваемых случаев эти алгебры являются групповыми алгебрами циклических групп или их фактор-алгебрами. Умножение элементов таких алгебр эквивалентно умножению в *полиномиальных кольцах* (или циклической свертке). Это позволяет воспользоваться известными быстрыми алгоритмами циклических сверток с минимальным числом умножений.

2. Конечномерные алгебры, ассоциированные с БА ДКП

В работе рассматриваются d-мерные алгебры A над R с базисом $\{\mathbf{1},\mathbf{e}_1,\ldots,\mathbf{e}_{d-1}\}$. Типичный элемент $\mathbf{a}\in\mathrm{A}$ записывается в виде

$$\mathbf{a}=\alpha_0\cdot\mathbf{1}+\sum_{j=1}^{d-1}\alpha_j\mathbf{e}_j\quad.$$

Сложение элементов выполняется покомпонентно, а умножение индуцируется заданными правилами умножения базисных элементов и определяет структуру алгебры на векторном пространстве P^d.

В работе рассматриваются следующие конечномерные алгебры.

1. Двумерная алгебра $\mathrm{A}_1^{(2)}$ с базисом $\{\mathbf{1},\mathbf{e}_1\}$ и правилами умножения базисных элементов:

$$\mathbf{1}^2=1\ ,\ \ \mathbf{e}_1^2=-1\ \cdot$$

(Алгебра C комплексных чисел).

2. Двумерная алгебра $\mathrm{A}_2^{(2)}$ с базисом $\{\mathbf{1},\mathbf{e}_1\}$ и правилами умножения базисных элементов:

$$\mathbf{1}^2=1\ ,\ \ \mathbf{e}_1^2=1\ \cdot$$

(Алгебра «двойных» чисел [18, 19], изоморфная прямой сумме $\mathrm{P}\oplus\mathrm{P}$.)

3. Трехмерная алгебра $\mathrm{A}^{(3)}$ с базисом $\{\mathbf{1},\mathbf{e}_1,\mathbf{e}_2\}$ и правилами умножения базисных элементов:

$$\mathbf{e}_1^2=\mathbf{e}_2\ ,\ \mathbf{e}_2^2=-\mathbf{e}_1\ ,\ \mathbf{e}_1\mathbf{e}_2=\mathbf{e}_2\mathbf{e}_1=-1\ .$$

4. Четырехмерная алгебра $\mathrm{A}_1^{(4)}$ с базисом $\{\mathbf{1},\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}$ и правилами умножения базисных элементов:

$$e_1^2=-e_2\ ,\ e_2^2=-1\ ,\ e_3^2=e_2\ ,\ e_1e_2=e_2e_1=e_3\ ,$$
$$e_2e_3=e_3e_2=-e_1\ ,\ e_1e_3=e_3e_1=1\ .$$

5. Четырехмерная алгебра $\mathrm{A}_2^{(4)}$ с базисом $\{\mathbf{1},\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}$ и правилами умножения базисных элементов:

$$e_1^2=e_2\ ,\ e_2^2=1\ ,\ e_3^2=e_2\ ,\ e_1e_2=e_2e_1=e_3\ ,$$
$$e_2e_3=e_3e_2=e_1\ ,\ e_1e_3=e_3e_1=1\ .$$

6. Четырехмерная алгебра $\mathrm{A}_3^{(4)}$ с базисом $\{\mathbf{1}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ и правилами умножения базисных элементов:

$$e_1^2 = -1\ ,\ e_2^2 = 1\ ,\ e_3^2 = -1\ ,$$

$$e_1e_2 = e_2e_1 = -e_3\ ,\ e_2e_3 = e_3e_2 = -e_1\ ,$$

$$e_1e_3 = e_3e_1 = e_2\ . \qquad (3)$$

Утверждение 1

а) Умножение элементов

$(\alpha_0 + \alpha_1\mathbf{e}_1)\ ,\ (\beta_0 + \beta_1\mathbf{e}_1)\ \in \mathrm{A}_1^{(2)}$

равносильно умножению полиномов

$(\alpha_0 + \alpha_1 t)(\beta_0 + \beta_1 t)\left(\mathrm{mod}\left(t^2 + 1\right)\right)$

и требует 3 умножений и 3 сложений [20].

б) Умножение элементов

$(\alpha_0 + \alpha_1\mathbf{e}_1)\ ,\ (\beta_0 + \beta_1\mathbf{e}_1)\ \in \mathrm{A}_2^{(2)}$

равносильно умножению полиномов

$(\alpha_0 + \alpha_1 t)(\beta_0 + \beta_1 t)\left(\mathrm{mod}\left(t^2 - 1\right)\right)$

и требует 2 умножений и 4 сложений [20].

в) Умножение элементов

$(\alpha_0 + \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2), (\beta_0 + \beta_1\mathbf{e}_1 + \beta_2\mathbf{e}_2) \in \mathrm{A}^{(3)}$

равносильно умножению полиномов

$\left(\alpha_0 + \alpha_1 t + \alpha_2 t^2\right)\left(\beta_0 + \beta_1 t + \beta_2 t^2\right)\left(\mathrm{mod}\left(t^3 + 1\right)\right)$

и требует 4 умножений и 14 сложений [20].

г) Умножение элементов

$(\alpha_0 + \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \alpha_3\mathbf{e}_3)\ ,$

$(\beta_0 + \beta_1\mathbf{e}_1 + \beta_2\mathbf{e}_2 + \beta_3\mathbf{e}_3)\ \in \mathrm{A}_1^{(4)}$

равносильно умножению полиномов

$\left(\alpha_0 + \alpha_1 t - \alpha_2 t^2 - \alpha_3 t^3\right)\left(\beta_0 + \beta_1 t - \beta_2 t^2 - \beta_3 t^3\right)\left(\mathrm{mod}\left(t^4 + 1\right)\right)$

и требует 9 умножений и 15 сложений [20].

д) Умножение элементов

$(\alpha_0 + \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \alpha_3\mathbf{e}_3)\ ,$

$(\beta_0 + \beta_1\mathbf{e}_1 + \beta_2\mathbf{e}_2 + \beta_3\mathbf{e}_3)\ \in \mathrm{A}_2^{(4)}$

равносильно умножению полиномов

$\left(\alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3\right)\left(\beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3\right)\left(\mathrm{mod}\left(t^4 - 1\right)\right)$

и требует 5 умножений и 15 сложений [20].

Утверждение 2

Алгебра $\mathrm{A}_3^{(4)}$ изоморфна прямой сумме $\mathrm{X} \oplus \mathrm{X}$:

$$\mathrm{C} \oplus \mathrm{C} = \left\{ \begin{array}{l} (z_1, z_2) : z_1 = a_1 + b_1 i_1, \\ z_2 = a_2 + b_2 i_2,\ i_1^2 = i_2^2 = -1, \\ a_1, a_2, b_1, b_2 \in \mathrm{R} \end{array} \right\}$$

Доказательство

Элементы

$\mathbf{E}_0 = (1,1)\ ,\ \mathbf{E}_1 = (i_1, -i_2)\ ,\ \mathbf{E}_2 = (-1,1)\ ,$

$\mathbf{E}_3 = (i_1, i_2)\ \in\ \mathrm{X} \oplus \mathrm{X}$

образуют базис алгебры $\mathrm{X} \oplus \mathrm{X}$ над R. Отображение φφ, определенное для базисных элементов алгебр $\mathrm{A}_3^{(4)}$ и $\mathrm{X} \oplus \mathrm{X}$ как

$\varphi : \mathbf{e}_1 \mapsto (i_1, -i_2)\ ,\quad \varphi : \mathbf{e}_2 \mapsto (-1,1)\ ,$

$\varphi : \mathbf{e}_3 \mapsto (i_1, i_2)\ ,\ \varphi : 1 \mapsto (1,1)$

продолжается R-линейно до изоморфизма соответствующих четырехмерных пространств и сохраняет равенства (3).

Линейный оператор L, определенный на пространстве алгебры $\mathrm{X} \oplus \mathrm{X}$ образами базисных элементов

$L(\mathbf{E}_0) = \frac{1}{2}(\mathbf{E}_0 - \mathbf{E}_2)\ ,\ L(\mathbf{E}_1) = \frac{1}{2}(\mathbf{E}_1 + \mathbf{E}_3)\ ,$

$L(\mathbf{E}_2) = \frac{1}{2}(\mathbf{E}_0 + \mathbf{E}_2)\ ,\ L(\mathbf{E}_3) = \frac{1}{2}(\mathbf{E}_3 - \mathbf{E}_1)\ ,$

преобразует базис $\{\mathbf{E}_0, \mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3\}$ в «стандартный» базис $\{\sigma_0, \sigma_1, \sigma_2, \sigma_3\}$ алгебры $\mathrm{X} \oplus \mathrm{X}$, рассматриваемой как четырехмерная R-алгебра:

$\sigma_0 = L(\mathbf{E}_0) = (1 + 0 \cdot i_1, 0 + 0 \cdot i_2)\ ,$

$\sigma_1 = L(\mathbf{E}_1) = (0 + i_1, 0 + 0 \cdot i_2)\ ,$

$\sigma_2 = L(\mathbf{E}_2) = (0 + 0 \cdot i_1, 1 + 0 \cdot i_2)\ ,$

$\sigma_3 = L(\mathbf{E}_3) = (0 + 0 \cdot i_1, 0 + i_2)\ .$

Следствие 1

Умножение постоянного элемента $\mathbf{a} = \alpha_0 + \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \alpha_3\mathbf{e}_3$ на вектор $\mathbf{b} = \beta_0 + \beta_1\mathbf{e}_1 + \beta_2\mathbf{e}_2 + \beta_3\mathbf{e}_3$ алгебры $\mathrm{A}_3^{(4)}$ требует 6 вещественных умножений и 10 вещественных сложений.

Доказательство

Сложность рассматриваемого умножения складывается из умножения двух пар комплексных чисел (элементов алгебры $\mathrm{X}\oplus\mathrm{X}$ в базисе $\{\sigma_0,\sigma_1,\sigma_2,\sigma_3\}$) и сложности преобразования элементов при замене базиса $\{\mathbf{E}_0,\mathbf{E}_1,\mathbf{E}_2,\mathbf{E}_3\}$ на базис $\{\sigma_0,\sigma_1,\sigma_2,\sigma_3\}$. Для умножения двух пар комплексных чисел достаточно 3+3 умножений и 3+3 сложений [20]. Для преобразования элементов переменного вектора достаточно 4 сложений.

3. Базовый пример: ДКП длины 8

В ненормализованной матричной форме ДКП (1) принимает вид $\mathbf{X}=\mathbf{F}\mathbf{x}$, где

$$\mathbf{X}^t=\left(X(0),\ldots,X(7)\right),$$

$$\mathbf{x}^t=\left(x(0),\ldots x(7)\right),$$

$\mathbf{F}$ - матрица ДКП, t - знак транспонирования.

После переупорядочивания компонент входного и выходного векторов

$$\mathbf{Y}^t=(Y(0),\ldots,Y(7))=$$
$$=(X(1),X(5),X(7),X(3),X(2),X(6),X(4),X(0))$$

$$\mathbf{y}^t=(y(0),\ldots,y(7))=$$
$$=(x(0),x(2),x(4),x(6),x(7),x(5),x(3),x(1))$$

матричное представление ДКП может быть записано в форме $\mathbf{Y}=\mathbf{T}\mathbf{y}$, где

$$\mathbf{T}=\left(\begin{array}{cc|cc|cc|cc}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1\\
g & -g & g & -g & g & -g & g & -g\\
\hline
f & -e & -f & e & f & -e & -f & e\\
e & f & -e & -f & e & f & -e & -f\\
\hline
a & c & -d & -b & -a & -c & d & b\\
c & d & -b & a & -c & -d & b & -a\\
d & b & a & c & -d & -b & -a & -c\\
b & -a & c & d & -b & a & -c & -d
\end{array}\right),$$

$$a=\cos\left(\frac{\pi}{16}\right),\ b=\cos\left(\frac{3\pi}{16}\right),\ c=\cos\left(\frac{5\pi}{16}\right),$$

$$d=\cos\left(\frac{7\pi}{16}\right),\ e=\cos\left(\frac{6\pi}{16}\right),\ f=\cos\left(\frac{2\pi}{16}\right),$$

$$g=\cos\left(\frac{4\pi}{16}\right).$$

Формирование из компонент вектора $\mathbf{y}$ вспомогательного массива:

$$z(4)=y(0)-y(4),\ z(5)=y(1)-y(5),$$

$$z(6)=y(2)-y(6),\ z(7)=y(3)-y(7),$$

$$z(2)=\left(y(0)+y(4)\right)-\left(y(2)+y(6)\right),$$

$$z(3)=\left(y(1)+y(5)\right)-\left(y(3)+y(7)\right),$$

$$z(6)=\left[(y(0)+y(4))+(y(2)+y(6))\right]-$$
$$-\left[(y(1)+y(5))+(y(3)+y(7))\right]$$

$$z(7)=\left[(y(0)+y(4))+(y(2)+y(6))\right]+$$
$$+\left[(y(1)+y(5))+(y(3)+y(7))\right] \qquad (4)$$

требует 14 операций вещественного сложения. После этого выполнение косинусного преобразования сводится к следующим матричным вычислениям:

$$\begin{pmatrix}Y(4)\\Y(5)\\Y(6)\\Y(7)\end{pmatrix}=\begin{pmatrix}a & c & -d & -b\\ c & d & -b & a\\ d & b & a & c\\ b & -a & c & d\end{pmatrix}\begin{pmatrix}z(4)\\z(5)\\z(6)\\z(7)\end{pmatrix}, \qquad (5)$$

$$\begin{pmatrix}Y(2)\\Y(3)\end{pmatrix}=\begin{pmatrix}f & -e\\ e & f\end{pmatrix}\begin{pmatrix}z(2)\\z(3)\end{pmatrix}, \qquad (6)$$

$$Y(0)=z(0),$$
$$Y(1)=gz(1) \qquad (7)$$

Утверждение 3

а) Вычисление матричного произведения (5) эквивалентно вычислению произведения элементов $\mathbf{s},\mathbf{p}\in\mathrm{A}_1^{(4)}$:

$$AX = (\alpha e_1 + \beta + \gamma e_2 + \delta e_3)(x + y e_1 + z e_2 + w e_3)$$

и, в соответствии с утверждением 1(г), требует 9 операций вещественного умножения и 15 операций вещественного сложения.

б) Вычисление матричного произведения (6) эквивалентно вычислению произведения элементов $\mathbf{q}, \mathbf{r} \in \mathrm{A}_1^{(2)}$:

$$\mathbf{qr} = (f + l\mathbf{e}_1)(z(4) + z(5)\mathbf{e}_1)$$

и, в соответствии с утверждением 1(а) требует 3 операций вещественного умножения и 3 операций вещественного сложения.

в) Вычисление по формуле (7) требует одной операции вещественного умножения.

Суммарная сложность алгоритма ДКП длины 8 с учетом формирования вспомогательных переменных $z_0,...,z_7$ составляет 9+3+1=13 операций умножения и 14+15+3=32 операции сложения.

Замечание 1

Структура рассмотренного алгоритма не зависит от конкретных значений параметров $a, b, ..., g$. Пусть

$$g' = e' = c' = 1, \; f' = f/e,$$

$$a' = a/c, \; d' = d/c, \; b' = b/c.$$

Тогда умножение в (7) становится тривиальным, в матричном произведении (6) остается два умножения. Вычисление правой части соотношения (5) требует 8 операций умножения. А умножения на g, e, c объединяются с нормализацией компонент косинусного спектра (с умножениями на λ_m) в (1).

Таким образом, предложенный алгоритм ДКП длины 8 требует 2+8=10 операций умножения и 32 операции сложения.

4. Алгоритмы ДКП длины 9, 10, 12, 15

После перестановки ряда строк и столбцов матрицы ДКП длины 9, 10, принимают вид

$$\mathbf{T}_9 = \left(\begin{array}{ccc:ccc|cc:c}
a & c & d & -d & -c & -a & b & -b & 0 \\
c & -d & a & -a & d & -c & -b & b & 0 \\
d & a & -c & c & -a & -d & -b & b & 0 \\
\hline
e & -f & -g & -g & -f & e & h & h & -1 \\
g & -e & f & f & -e & g & -h & -h & 1 \\
f & g & -e & -e & g & f & -h & -h & 1 \\
\hline
b & -b & -b & b & b & -b & 0 & 0 & 0 \\
\hdashline
h & h & h & h & h & h & -1 & -1 & -1 \\
\hdashline
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}\right),$$

$$\mathbf{T}_{10} = \left(\begin{array}{cc|cc|cc|cc|cc}
a & b & c & d & -d & -c & -b & -a & q & -q \\
b & d & -a & -c & c & a & -d & -b & -q & q \\
c & -a & d & -b & b & -d & a & -c & q & -q \\
d & -c & -b & a & -a & b & c & -d & q & -q \\
\hline
e & f & -f & -e & -e & -f & f & e & 0 & 0 \\
f & -e & e & -f & -f & e & -e & f & 0 & 0 \\
\hline
g & -h & -h & g & g & -h & -h & g & -1 & -1 \\
h & -g & -g & h & h & -g & -g & h & 1 & 1 \\
\hline
q & -q & q & q & -q & -q & q & -q & -q & q \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}\right).$$

Утверждение 4

Матричные умножения

$$\begin{pmatrix} a & c & d \\ c & -d & a \\ d & a & -c \end{pmatrix} \begin{pmatrix} z(0) \\ z(1) \\ z(2) \end{pmatrix}, \begin{pmatrix} e & -f & -g \\ g & -e & f \\ f & g & -e \end{pmatrix} \begin{pmatrix} z(5) \\ z(6) \\ z(7) \end{pmatrix},$$

эквивалентны умножению элементов алгебры $\mathrm{A}^{(3)}$:

$$(-a\mathbf{e}_2 - c\mathbf{e}_1 + d)(z(0)\mathbf{e}_1 + z(1)\mathbf{e}_2 + z(2))$$

и

$$(f\mathbf{e}_2 - g\mathbf{e}_1 - e)(z(6)\mathbf{e}_2 - z(5)\mathbf{e}_1 + z(7))$$

соответственно и требуют согласно утверждению 1(в) 4 вещественных умножений и 14 вещественных сложений каждое.

Следствие 2

ДКП длины 9 посредством умножения на матрицу $\mathbf{T}_9$ выполняется за 8 умножений и 44 сложения (то есть требует менее одного умножения и около пяти сложений на отсчет).

Утверждение 5

а) Матричное умножение

$$\mathbf{Ty} = \begin{pmatrix} \alpha & \beta & \gamma & \delta \\ \beta & \delta & -\alpha & -\gamma \\ \gamma & -\alpha & \delta & -\beta \\ \delta & -\gamma & -\beta & \alpha \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} \alpha x + \beta y + \gamma z + \delta w \\ \beta x + \delta y - \alpha z - \gamma w \\ \gamma x - \alpha y + \delta z - \beta w \\ \delta x - \gamma y - \beta z + \alpha w \end{pmatrix}$$

эквивалентно вычислению произведения элементов алгебры $A_2^{(4)}$

$$AX = (\alpha e_2 + \beta e_1 - \gamma e_3 + \delta)(x e_1 + y e_2 - z + w e_3).$$

и требует согласно утверждению 1(д) 5 вещественных умножений и 15 вещественных сложений.

б) Вычисление матричного произведения

$$\begin{pmatrix} l & f \\ f & -l \end{pmatrix} \begin{pmatrix} z(4) \\ z(5) \end{pmatrix}$$

эквивалентно вычислению произведения элементов алгебры $A_1^{(2)}$:

$$AX = (\alpha e_1 + \beta + \gamma e_2 + \delta e_3)(x + y e_1 + z e_2 + w e_3)$$

и, в соответствии с утверждением 1(а) требует 3 операций вещественного умножения и 3 операций вещественного сложения.

в) Вычисление матричного произведения

$$\begin{pmatrix} g & -h \\ h & -g \end{pmatrix} \begin{pmatrix} z(6) \\ z(7) \end{pmatrix}$$

эквивалентно вычислению произведения элементов алгебры $A_2^{(2)}$:

$$AX = (\alpha e_1 + \beta + \gamma e_2 + \delta e_3)(x + y e_1 + z e_2 + w e_3)$$

и, в соответствии с утверждением 1(б) требует 2 операций вещественного умножения и 4 операций вещественного сложения.

Следствие 3. ДКП длины 10 посредством умножения на матрицу $\mathbf{T}_{10}$ выполняется за 9 умножений и 43 сложения.

Аналогичные утверждения справедливы для N=12, 15.

Утверждение 6. Умножение матрицы ДКП длины 12 на входной вектор эквивалентно

а) умножению переменного элемента алгебры $A_3^{(4)}$ на постоянный элемент этой же алгебры;

б) умножению переменного элемента алгебры $A_1^{(2)}$ на постоянный элемент этой же алгебры;

в) умножению переменного элемента алгебры $A_2^{(2)}$ на постоянный элемент этой же алгебры;

г) дополнительным умножениям констант на переменные и вспомогательным сложениям.

Следствие 4. ДКП длины 12 посредством умножения на матрицу $\mathbf{T}_{12}$ выполняется за 13 умножений и 55 сложений.

Утверждение 7. Умножение матрицы ДКП длины 15 на входной вектор эквивалентно

а) двум умножениям переменного элемента алгебры $A_2^{(4)}$ на постоянный элемент этой же алгебры;

б) умножению переменного элемента алгебры $A_1^{(2)}$ на постоянный элемент этой же алгебры;

Таблица 1. Количество операций, необходимых для вычисления ДКП

N	Предложенные алгоритмы		Алгоритм работы [21]	
	*	+	*	+
8	10	32	12	29
9	8	44	11	44
10	9	43	15	36
12	13	51	20	43
15	21	82	35	89

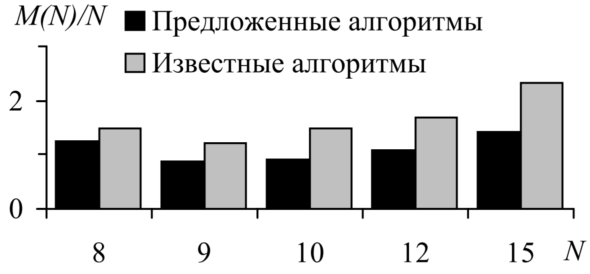


Рис. 1. Удельная мультипликативная сложность алгоритмов

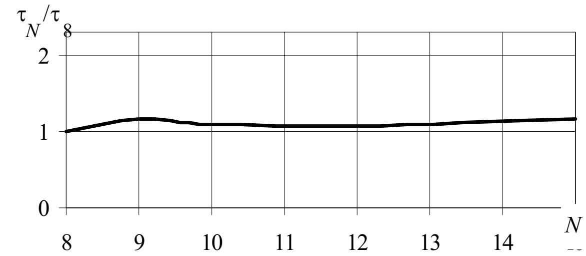


Рис. 2. Относительное время обработки изображения 1024π1024 пиксела блочным ДКП

в) умножению переменного элемента алгебры $\mathrm{A}_2^{(2)}$ на постоянный элемент этой же алгебры;

г) дополнительным умножениям констант на переменные и вспомогательным сложениям.

Следствие 5. ДКП длины 15 посредством умножения на матрицу $\mathbf{T}_{15}$ выполняется за 24 умножений и 83 сложения.

В качестве основы для сравнительного анализа вычислительной сложности синтезированных алгоритмов был использован алгоритм работы [21] синтезированный для ДКП произвольных длин, оценки сложности которого при $N = 2^k$ совпадают с оценками сложности лучших из известных алгоритмов ДКП [8-10].

В таблице 1 приведено количество операций необходимых для вычисления ДКП предложенным алгоритмом и известным способом. На рис.1 приводится зависимость удельной мультипликативной сложности алгоритмов от длины преобразования.

5. Результаты экспериментов

Целью экспериментального исследования было выявление зависимости времени τ_N обработки изображения блочным ДКП от размера квадратного блока *N*.

На рис. 2 показана относительная характеристика τ_N/τ_8, где τ_8 - это время обработки изображения блоками 8×8 ($\tau_N/\tau_8 = 1$ при *N*=8), размер изображения - 1024×1024 пиксела.

Полученные результаты позволяют утверждать, что предложенные алгоритмы ДКП гарантируют скорость обработки изображения близкую к скорости обработки лучшим из известных алгоритмов ДКП длины 8. Время обработки практически не возрастает с ростом *N*.

Заключение

В работе доказано, что синтезированные алгоритмы ДКП обладают минимальной мультипликативной сложностью. Методика синтеза таких алгоритмов может быть использована для построения одномерных ДКП ряда других длин, а также, с соответствующими изменениями, для синтеза алгоритмов двумерных ДКП. Разработанные высокоскоростные алгоритмы могут быть использованы при разработке новых быстродействующих адаптивных алгоритмов блочного кодирования, учитывающих спектральную неоднородность реальных изображений.

СПИСОК ЛИТЕРАТУРЫ

1. *Ахмед Н., Рао К. Р.* Ортогональные преобразования при обработке цифровых сигналов. Пер. с англ. М.: Связь, 248 с., 1980.
2. *Rao K. R., Yip P.* Discrete Cosine Transform. Academic Press, San Diego, 1990.
3. *Wallace G. K.* The JPEG still picture compression standard. //Communications of the ACM, Vol. 34, No 4, pp. 31-44, 1991.
4. *De Natale F. G. B., Desoli G. S., Giusto D. D., Vernazza G.* Adaptive DCT for image-data compression. //Eur. Trans. Telecommun. and Relat. Technol., Vol. 3, No 4, pp. 359-366, 1992
5. *Jeong J., Jo J. M.* Adaptive Huffman coding of 2-D DCT coefficients for image sequence

compression. //Signal Processing: Image Communication, Vol. 7, Issue 1, pp. 1-11, 1995

6. *Krupiczka A.* Interblock variance as a segmentation criterion in image coding. // Mashine Graphics and Vision, Vol. 5, Nos 1/2, pp. 229-235, 1996
7. *Sikora T.* Low complexity shape-adaptive DCT for coding of arbitrarily shaped image segments. // Signal Processing: Image Communication, Special Issue on Coding Techniques for Very Low Bit-rate Video, Vol. 7, Issue 4-6, pp. 381-395, 1995
8. *Chan-Wan Y.-H., Siu C.* On the realization of discrete cosine transform using the distributed arithmetic. //IEEE Transactions on Circuits and Systems - I: Fundamental Theory and Applications, Vol. 39, No 9, pp. 705-712, 1992
9. *Hou H. S., Tretter D. K.* Interesting properties of the discrete cosine transform. //J. Visual Commun. and Image Represent., Vol. 3, No 1, pp. 73-83, 1992
10. *Suheiro N., Hatori M.* Fast algorithms for the DFT and other sinusoidal transforms. //IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-34, No 6, pp. 642-644, 1986
11. *Власенко В. А., Лаппа Ю. М., Ярославский Л. П.* Методы синтеза быстрых алгоритмов свертки и спектрального анализа сигналов. М.: Наука, 1990. 160 с.
12. *Duhamel P.* Implementation of split-radix FFT algorithms for complex, real, and real-symmetric data. //IEEE Trans. Accoust., Speech, and Signal Process., 1984, Vol.32, No 4, pp.750-761.
13. *Heideman M. T.* Computation of an odd-length DCT from a real-valued DFT of the same length. //IEEE Trans. Signal Process., Vol. 40, No 1, pp.54-61, 1992
14. *Чернов В. М.* Алгоритмы двумерных дискретных ортогональных преобразований, реализуемые в кодах Гамильтона-Эйзенштейна. //Проблемы передачи информации, том 31, № 3, с.38-46, 1995.
15. *Chichyeva M. A., Chernov V. M.* «One-step» short-length DCT algorithms with data representation in the direct sum of associative algebras. //Proceedings CAIP'97, Springer, LNCS 1296, pp.590-596, 1997
16. *Feig, E. Ben-Or M.* On algebras related to the discrete cosine transform. //Linear Algebra and Its Applications, Vol. 266, pp. 81-106, 1997
17. *Bazensky, G. Tasche M.* Fast polynomial multiplication and convolutions related to the discrete cosine transform. //Linear Algebra and Its Applications, Vol. 252, Issue 1-3, pp. 1-25, 1997
18. *Jacobson N.* Structure and representation of Jordan algebras. Providence, R.I., 1968
19. *Schafer R.D.* An introduction to nonassociative algebras. London: Academic Press, 1966
20. *Nussbaumer H. J.* Fast Fourier transform and convolution algorithms. Springer Verlag, 1971.
21. *Chan S.-C., Ho K.-L.* Fast algorithms for computing the discrete cosine transform. //IEEE Trans. on Circuits and Systems, Vol. 39, No 3, pp.185-190, 1992.

[1] Работа выполнена при финансовой поддержке Российского Фонда Фундаментальных Исследований, (Грант № 97-01-00900)

FAST ALGORITHMS FOR DISCRETE COSINE TRANSFORMING OF SHORT LENGTHS WITH MINIMAL COMPUTATIONAL COMPLEXITY

© 1999 V.M. Chernov, M.A. Chichyeva

Image Processing System Institute of Russian Academy of Sciences, Samara

This paper considers fast algorithms for the discrete cosine transforming (DCT) of short lengths with minimal computational complexity. Decreasing of the computational complexity is achieved due to new approach of synthesis of the short lengths DCT algorithms. This approach is connected with interpretation of DCT calculation as operations within associated algebraic structures. The application of developed algorithms in block coding method is researched.

УДК 535.42

СПИРАЛЬНЫЕ ПУЧКИ СВЕТА - НОВОЕ НАПРАВЛЕНИЕ КОГЕРЕНТНОЙ ОПТИКИ

© 1999 Е. Г. Абрамочкин, В.Г. Волостников

Самарский филиал Физического института им. П.Н. Лебедева РАН

Теоретически найден и экспериментально реализован новый класс лазерных пучков, названных спиральными, которые сохраняют структуру своей интенсивности при распространении и фокусировке с точностью до масштаба и вращения. Получены спиральные пучки, распределение интенсивности которых имеет форму произвольной плоской линии. Описаны методы синтеза таких пучков и приведены результаты экспериментов.

Введение

Как известно, распространение светового поля, в частности лазерного пучка, представляет собой волновое явление и, как всякий колебательный процесс, характеризуется распределением амплитуды и фазы.

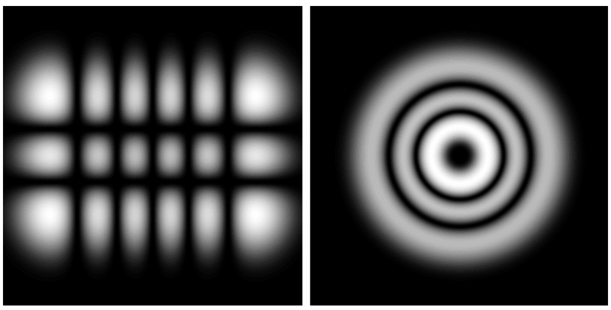
Если распределения амплитуды и фазы поля заданы в некоторой плоскости, то последующая эволюция поля при распространении описывается тем или иным дифференциальным уравнением. Отсюда следует, что световое поле при распространении, вообще говоря, претерпевает количественные и качественные изменения [1].

Однако, с открытием лазеров и появлением когерентной оптики, описывающей распространение лазерных пучков, было теоретически и экспериментально показано, что лазер может излучать световые пучки, которые самосогласованы таким образом, что сохраняют свою структуру при распространении и фокусировке с точностью до масштаба [2]. Такие пучки являются собственными колебаниями лазерных резонаторов (или модами), имеют жестко заданную форму и описываются двумя семействами специальных функций с различными типами симметрии: пучки Эрмита-Гаусса и Лагерра-Гаусса, при этом низший тип колебаний в этих семействах совпадает и является известной двумерной функцией Гаусса. На рис.1 приведены примеры поперечных распределений интенсивности высших типов пучков [3]. Структура этих пучков сохраняется при распространении и фокусировке и может ассоциироваться с однородными деформациями растяжения-сжатия: сходящиеся и расходящиеся пучки. Закономерно поставить вопрос: имеется ли некая оптическая аналогия деформации кручения для пучков с неоднородной расходимостью? Как показано в [4], поток световой энергии состоит в общем случае из двух компонент: потенциальной и вихревой. В определенном смысле, первая компонента соответствует деформациям растяжения-сжатия, а вторая – деформациям кручения.

Принимая во внимание вихревую компоненту потока световой энергии, можно расширить понятие структурной устойчивости световых полей, а именно, поставить следующую задачу: существуют ли лазерные пучки, сохраняющие свою структуру при распространении и фокусировке с точностью до масштаба и вращения?

Спиральные пучки света

В работе [7] найден и полностью описан новый класс лазерных пучков, названных спиральными, которые сохраняют свою



a b

Рис.1. Примеры распределений интенсивности пучка Эрмита-Гаусса (a) и пучка Лагерра-Гаусса (b) структуру при распространении и фокусировке и могут иметь различные параметры вращения.

Известные упомянутые выше типы лазерных мод Эрмита-Гаусса и Лагерра-Гаусса являются частными случаями спиральных пучков с нулевым вращением. Отличие найденных пучков от известных в том, что они обладают двумя принципиально новыми свойствами.

Во-первых, сохраняя форму при распространении и фокусировке, они могут иметь весьма разнообразную структуру распределения интенсивности. В частности, в семействе спиральных пучков теоретически найдены и на отдельных примерах реализованы экспериментально спиральные пучки в форме произвольных кривых или их совокупности [10]. Комплексная амплитуда найденных пучков в виде кривой, заданной в комплексной параметрической форме $z=z(t)=x(t)+iy(t)$, tO$[0,T]$, определяется выражением

$$\mathrm{S}\left(z,\bar{z}\middle|\zeta(t),\ t\in[0,T]\right)= \\ =\exp\left(-\frac{z\bar{z}}{\rho^2}\right)\int_0^T \exp\left[\begin{array}{l}-\dfrac{\zeta(t)\bar{\zeta}(t)}{\rho^2}+\dfrac{2z\bar{\zeta}(t)}{\rho^2}+\\ +\dfrac{1}{\rho^2}\displaystyle\int_0^t\left(\bar{\zeta}(\tau)\zeta'(\tau)-\zeta(\tau)\bar{\zeta}'(\tau)\right)d\tau\end{array}\right]\left|\zeta'(t)\right|dt. \tag{1}$$

Здесь r – гауссов параметр. Кривая $z(t)$ названа порождающей для спирального пучка (1). Примеры пучков , построенных по формуле (1), приведены на рис.2.

Отметим еще раз, что в отличие от световых полей с заданным распределением интенсивности, формируемых ранее известными методами, эти спиральные пучки сохраняют свою структуру в любой плоскости наблюдения и фокусировки. Таким образом, данное свойство спиральных пучков позволяет весьма гибко менять их форму при сохранении структурной устойчивости, что представляет существенный интерес для лазерной медицины и технологии.

Во-вторых, вихревой характер распространения световой энергии в пучках обуславливает то, что пучки обладают существенно ненулевым угловым моментом количества движения [8]. Это проявляется в том, что микроскопические объекты размерами в десятки микрон (например, живые клетки), помещенные в область фокусировки такого пучка могут приводиться во вращение вокруг своего центра инерции, удерживаться в заданной области пространства, подвергаться неоднородным заданным деформациям и т.п.

Эти два свойства дают возможность создания в области фокусировки заданных микрораспределений интенсивности и углового

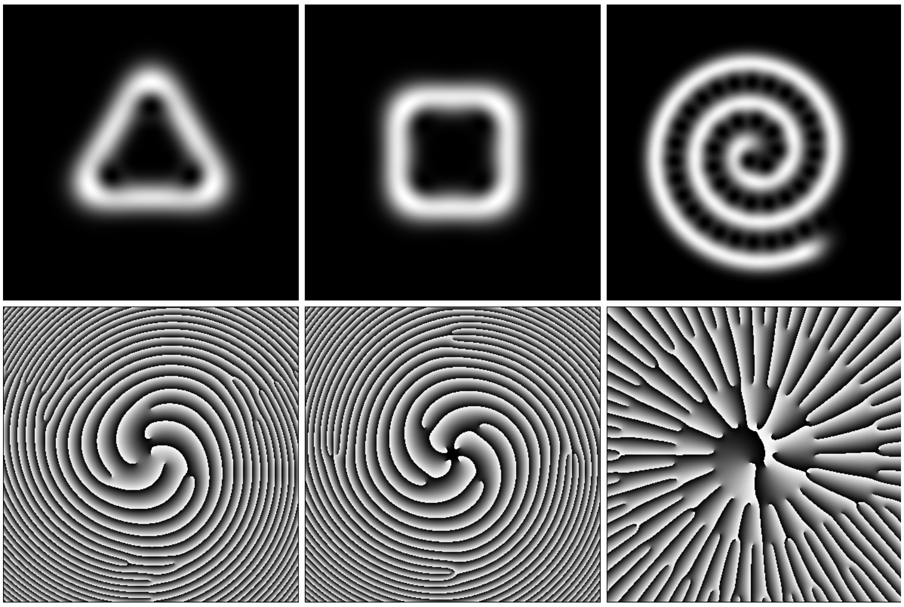


Рис.2. Примеры спиральных пучков в виде кривых: треугольник, квадрат и спираль Архимеда (интенсивности - верхний ряд, фазы - нижний ряд). Белый цвет соответствует фазе 2р, черный - нулевой. Концевые точки изофазных линий показывают местоположение изолированных нулей интенсивности, называемых также оптическими вихрями

момента, что представляет принципиально новый инструмент для бесконтактного манипулирования микрообъектами в электронике и микробиологии [5].

При исследовании свойств спиральных пучков для замкнутых порождающих кривых было установлено, что такие пучки проявляют характерные свойства квантования: во-первых, распределение интенсивности таких пучков претерпевает радикальное изменение при преобразовании подобия кривой $\mathrm{z}(t) \rightarrow \mathrm{nz}(t)$ и обладает формой кривой лишь при определенных дискретных значениях nn; во-вторых, для этих же дискретных значений nn распределения интенсивности пучков для кривых $\mathrm{nz}(t+a)$ при различных a одинаковы; в-третьих, площадь S под кривой $\mathrm{z}(t)$ связана с гауссовым параметром r такого пучка соотношением:

$$S==\tfrac{1}{2}\mathrm{pr}^2N, \quad \text{где } N=1,2,\ldots\ldots \qquad (2)$$

Пучки, порождающие кривые которых удовлетворяют равенству (2), названы N-квантованными.

Следует кратко отметить и квантово-механический аспект полученных результатов. Он обусловлен тем, что эволюция лазерных пучков при распространении описывается параболическим дифференциальным уравнением, полностью аналогичным уравнению Шредингера, описывающим эволюцию волновых функций квантовых систем во времени. При этом структурно устойчивым лазерным пучкам соответствуют в квантовой механике стационарные решения уравнения Шредингера с некоторым гамильтонианом. В работе [10] показано, что найденным спиральным пучкам соответствуют специфические стационарные состояния квантовой частицы в однородном магнитном поле, причем для квантово-механической аналогии – состояний заряженной частицы в однородном магнитном поле – условие (2) соответствует квантованному магнитному потоку через контур $\mathrm{z}(t)$: $\Phi=(2\mathrm{p}\hbar c/|e|)N$. Данная аналогия представляет интерес в приложениях к теории квантового эффекта Холла [6].

Результаты экспериментов

Эспериментальная реализация спиральных пучков осуществлялась различными ме-

Рис.3. Экспериментальное распределение интенсивности спирального пучка в форме границы треугольника (см. теоретическое распределение на рис.2)

тодами. Во-первых, непосредственно с помощью амплитудно-фазовых масок. Амплитудная маска и фотошаблон для фазовой маски изготавливались на фотоплоттере (разрешение 1024×1024, размер 10×10 мм^2). Фотошаблон использовался для изготовления фазового элемента на дихромированной желатине. Комбинация амплитудной и фазовой масок давала требуемый амплитудно-фазовый элемент, который освещался пучком гелий-неонового лазера. Пример экспериментальной реализации пучка в виде границы треугольника показан на рис.3. Пучок сохранял свою структуру при фокусировке и поворачивался на 90° при распространении из области перетяжки в дальнюю зону.

Другой, менее очевидный метод синтеза таких пучков основан на результатах работы [9], где, в частности, показано, что астигматическим воздействием можно осуществить преобразование пучков Лагерра–Гаусса в пучки Эрмита–Гаусса и наоборот. Это преобразование может быть обобщено следующим образом (см. [10]):

$$\iint\limits_{\mathrm{R}^2} \exp\left(-\mathrm{i}(x\xi + y\eta) + \frac{2\mathrm{i}\xi\eta}{\rho^2}\right)\mathrm{S}(\xi + \mathrm{i}\eta, \xi - \mathrm{i}\eta|\zeta)\,d\xi d\eta = \frac{\pi\rho^2}{\sqrt{2}}\exp\left(-\frac{\mathrm{i}\rho^2 xy}{4} - \frac{\rho^2 x^2}{8}\right)h(\rho y|\zeta),$$

где

$h(\mathrm{r}y|\mathrm{z})=$

$$= \exp\left(-\frac{\rho^2 y^2}{8}\right)\int\limits_0^T \exp\left[\begin{array}{l} -\dfrac{\bar{\zeta}^2(t)}{\rho^2} - \dfrac{\zeta(t)\bar{\zeta}(t)}{\rho^2} + y\bar{\zeta}(t) + \\ + \dfrac{1}{\rho^2}\int\limits_0^t \left(\bar{\zeta}\zeta' - \zeta\bar{\zeta}'\right)d\tau \end{array}\right]|\zeta'(t)|\,dt.$$

Таким образом, задача формирования спирального пучка посредством астигматических преобразований сводится к синтезу одномерного по структуре поля $h(ry|z)$. На рис.4 показано одномерное поле $h(ry|z)$ для кривой в виде границы треугольника и схема синтеза соответствующего спирального пучка. Одномерная структура поля позволяет в полной мере использовать возможности микролитографии, поэтому данный способ может быть технологически более предпочтителен, чем прямой метод амплитудно-фазовой маски.

С помощью предложенного метода были экспериментально реализованы пучки различного вида, сохраняющие свою структуру при распространении и фокусировке. На рис.5 показана экспериментальная реализация спирального пучка в виде регулярной решетки нулей, когда в качестве одномерной структуры была использована фазовая мультиплицирующая дифракционная решетка Даммана. Волновые поля такого вида (т.н. решетки вихрей) естественным образом возникают в квантово-механических задачах, связанных со сверхпроводимостью, сверхтекучестью, квантовым эффектом Холла [12].

Исследования спиральных пучков показали, что они являются собственными колебаниями специфических резонаторов с вращением поля. Был реализован аргоновый лазер с таким резонатором и впервые полу-

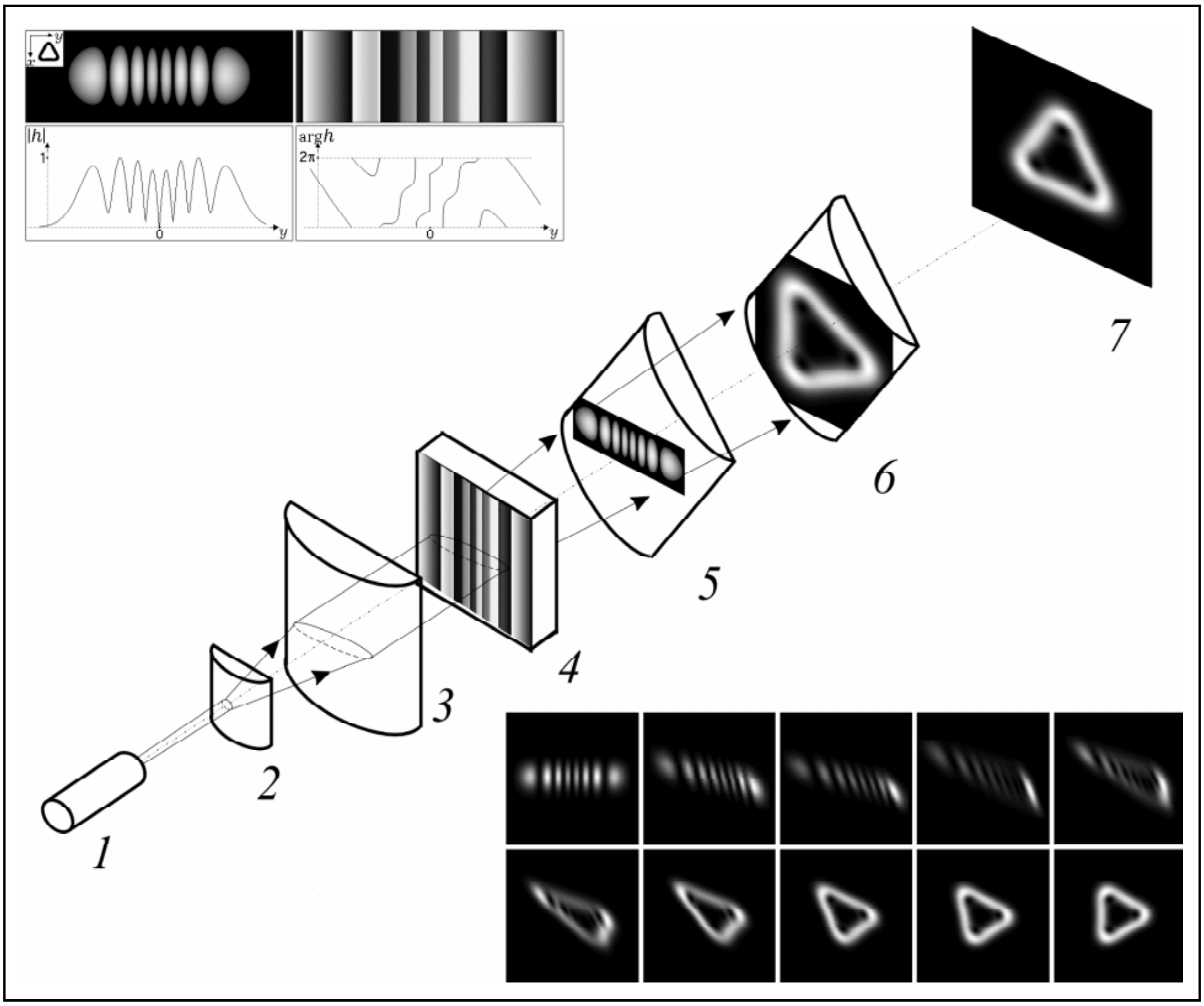
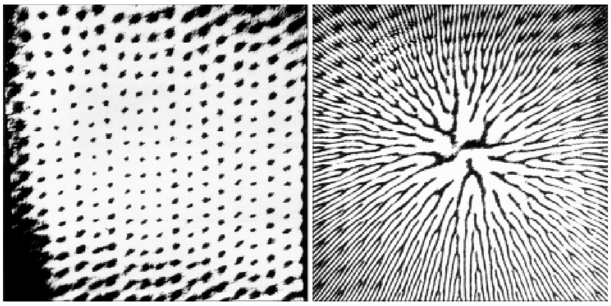


Рис.4. Оптическая схема синтеза спиральных пучков посредством одномерных по структуре оптических элементов (1 - лазер, 2,3,5,6 - цилиндрические линзы, 4 - одномерный оптический элемент, 7 - экран). В верхней части рисунка показана структура амплитудно-фазового элемента для спирального пучка в форме границы треугольника. Полутоновые изображения представляют собой амплитуду и фазу распределения $\exp(-r^2x^2/8)h(ry|z)$; графики соответствуют амплитуде и фазе одномерных распределений $h(ry|z)$. В нижней части рисунка показана динамика изменения светового поля в промежутке между линзами 5,6. За линзой 6 пучок сохраняет свою структуру и вращается при распространении

чены экспериментально спиральные пучки с различными параметрами вращения [11]. Это показывает принципиальную возможность создания лазеров, непосредственным результатом генерации которых без дополнительной нестандартной оптики будут пучки с заданными свойствами.

Использование фазовой структуры спиральных пучков с заданным распределением интенсивности как базовой дает новый подход и к известной задаче синтеза чисто фазовых элементов для фокусировки в кривые [13]. Обычно базовые решения ищутся на основе метода стационарной фазы. Можно показать (см., например, [7]), что решения на базе спиральных пучков невозможно получить таким образом. Пример решения задачи с использованием структуры спиральных пучков показан на рис.6 (амплитудная маска для фазового элемента изготовлена на фотоплоттере в Институте систем обработки изображений РАН, фазовый элемент на желатине сделан Н.Н.Лосевским). Характерной чертой получаемых решений является их однотипный топологический характер и наличие существенного количества фазовых сингулярностей (оптических вихрей), что указывает на их родство со спиральными пучками (ср., например, рис.6а и соответствующее фазовое распределение на рис.2).

Таким образом, создана теоретическая и экспериментальная основа принципиально новых возможностей целенаправленного “конструирования” лазерного излучения. На



a b

Рис.5. Экспериментальная реализация спирального пучка в форме решетки нулей: интенсивность (a) и результат интерференции между спиральным и опорным пучками (b)

рис.7 представлены примеры спиральных пучков сложной формы, которые, несмотря на свой «рукотворный» вид, являются такими же естественными физико-математическими объектами как и обычные лазерные пучки, являются точными решениями уравнения Шредингера, могут быть собственными колебаниями соответствующих лазерных резонаторов и сохраняют свою структуру при распространении и фокусировке.

Заключение

Основными результатами проведенного исследования являются следующие:

· Найдены новые лазерные пучки, названные спиральными, которые сохраняют структуру своей интенсивности при распространении и фокусировке с точностью до масштаба и вращения. Показано, что данные пучки могут быть модами специфических резо-

a b c

Рис.6. Фазовый элемент для фокусировки пучка с равномерной засветкой круговой апертуры в границу квадрата (a), теоретическое (b) и экспериментальное (c) распределения интенсивности в фокусе. Некоторое различие между теорией и экспериментом обусловлено погрешностями передачи контраста амплитудной маски и нелинейностями при формировании фазового рельефа слоем желатины

Рис.7. Спиральные пучки сложной структуры (интенсивности и фазы)

наторов с вращением поля. Экспериментально реализован лазер с таким резонатором и получены пучки с различными параметрами вращения при распространении. Выявлено соответствие спиральных пучков волновым функциям заряженной частицы в однородном магнитном поле.

· Теоретически обоснована и экспериментально показана возможность синтеза световых полей с заданными пространственными характеристиками и сохраняющих свою структуру при распространении и фокусировке. Найден класс лазерных пучков, распределение интенсивности которых имеет форму произвольной плоской линии. Разработаны основы оптики таких пучков.

· На основе полученных закономерностей преобразования пучков Лагерра-Гаусса в пучки Эрмита-Гаусса разработан метод синтеза спиральных пучков посредством одномерных амплитудно-фазовых элементов.

Данная работа частично поддержана Российским фондом фундаментальных исследований (грант 99-02-16513).

СПИСОК ЛИТЕРАТУРЫ

1. *М.Борн*, *Э.Вольф*. Основы оптики. М., Наука, 1973.
2. *H.Kogelnik*, *T.Li*. Laser beams and resonators // Applied Optics. 1966. Vol. 5. P. 1550.
3. *Ю.А.Ананьев*. Оптические резонаторы и лазерные пучки. М., Наука, 1990.
4. *E.Abramochkin*, *V.Volostnikov*. Relationship between two-dimensional intensity and phase in a Fresnel diffraction zone // Optics Communications. 1989. vol. 74. P. 144.
5. *M.Padgett*, *L.Allen*. Optical tweezers and spanners // Physics World. September 1997. P.35.
6. *В.А.Гейлер*. Двумерный оператор Шредингера с однородным магнитным полем и его возмущения периодическими потенциалами нулевого радиуса // Алгебра и анализ. 1991. Т. 3. Вып. 3. С. 2.
7. *E.Abramochkin*, *V.Volostnikov*. Spiral-type beams // Optics Communications. 1993. Vol. 102. P. 336.
8. *E.Abramochkin*, *V.Volostnikov*. Structurally stable singular wavefields // International Conference on Singular Optics. Proceedings of SPIE. 1998. Vol. 3487. P. 20.
9. *E.Abramochkin*, *V.Volostnikov*. Beam transformations and nontransformed beams // Optics Communications. 1991. Vol. 83. P. 123.
10. *E.Abramochkin*, *V.Volostnikov*. Spiral-type beams: optical and quantum aspects // Optics Communications. 1996. Vol. 125. P. 302.
11. *E.Abramochkin*, *N.Losevsky*, *V.Volostnikov*. Generation of spiral-type laser beams // Optics Communications. 1997. Vol. 141. P. 59.
12. *I.Dana*, *I.Freund*. Vortex-lattice wave fields // Optics Communications. 1997. Vol. 136. P. 93.
13. *А.В.Гончарский, В.В.Попов, В.В.Степанов.* Введение в компьютерную оптику. М., МГУ, 1991.

LIGHT BEAMS OF SPIRAL TYPE: A NEW FIELD IN COHERENT OPTICS

© 1999 E.G. Abramochkin, V.G. Volostnikov

Samara Branch of Physics Institute named for P.N. Lebedev of Russian Academy of Sciences

New class of laser beams, which are called spiral type beams, was theoretically researched and realized experimentally. These expanding beams have predetermined intensity distribution structure. The structure is accurate within scale and rotation. The spiral-type beams with intensity distribution in the form of arbitrary line were obtained. Several methods of spiral-type beam synthesis are described in the paper. Experimental results of transformation of Gaussian beam into a triangle-line laser beam are also presented.