

**БЫСТРЫЙ СПОСОБ ВЫЧИСЛЕНИЯ ИНТЕГРАЛА РЕЛЕЯ-ЗОММЕРФЕЛЬДА ПЕРВОГО ТИПА**

Андрей Владимирович Устинов (ведущий программист; e-mail: andr@smr.ru)

Учреждение Российской академии наук Институт систем обработки изображений РАН

**Аннотация**

В данной статье описаны приёмы ускорения вычисления интеграла Кирхгофа за счёт использования аналитических преобразований и эффективной программной реализации. Показаны последовательные шаги увеличения быстродействия после каждого преобразования. Получены условия применения различных приближений. Отметим, что после применения всех преобразований время выполнения программы слабо зависит от расстояния между входной и выходной плоскостями.

**Ключевые слова:** интеграл Кирхгофа, параксиальность, квадратурная формула, асимптотическое разложение.

**Введение**

Интеграл типа Кирхгофа (в зависимости от граничных условий различают дифракционные интегралы Кирхгофа или Релея-Зоммерфельда) в скалярной или векторной форме позволяет достаточно точно описывать дифракцию световых полей в ближней зоне ( $\ll$  длины волны) на малых фазовых объектах [1], на малых отверстиях в металлическом и черном экране [2].

Поэтому важной задачей является вычисление интеграла Кирхгофа за приемлемое время в тех случаях, когда параксиальное приближение Френеля неприменимо. На малых расстояниях допускается применение метода распространения плоских волн, который имеет большую скорость вычислений при использовании быстрого преобразования Фурье (БПФ), но он имеет определённые ограничения:

1. После того как выбрана сетка дискретизации входного поля, область вычислений фиксирована и в пространстве координат, и в пространстве частот. Это ограничивает область, в которой может быть вычислено выходное поле.

2. Входное поле должно убывать до нуля при приближении к краям области вычислений.

3. Наиболее важно, что при увеличении расстояния начинает появляться эффект перекрытия спектров. Это в принципе можно преодолеть, применяя в пространстве частот более густую сетку и интерполируя значения спектра Фурье входного поля на новой сетке. Это, однако, довольно громоздко и будет приводить к неприемлемому увеличению времени вычисления и объёма используемой памяти.

Первые два ограничения присущи и приближению Френеля, если оно реализовано с использованием БПФ.

Кроме того, область применимости метода распространения плоских волн не всегда покрывает область неприменимости приближения Френеля. Поэтому в работе [3] предложен метод, который работоспособен при любых расстояниях, но, как будет сказано ниже, его быстродействие недостаточно велико. В данной статье описывается более эффективный способ вычисления интеграла Кирхгофа.

**1. Общие соотношения**

Интеграл Кирхгофа выражает решение дифракционных задач при помощи теоремы [4], согласно которой амплитуда электромагнитного поля  $U(x, y, z)$  на выходе выражается через её значение  $U_0(x, y, 0)$  на входе:

$$U(x, y, z) = -\frac{ikz}{2\pi} \iint_S U_0(u, v, 0) \frac{e^{ikR}}{R^2} \left(1 + \frac{i}{kR}\right) dudv, \quad (1)$$

где  $R = \sqrt{(u-x)^2 + (v-y)^2 + z^2}$  - расстояние между точками на входной и выходной плоскостях,  $k = 2\pi/\lambda$  - волновое число,  $\lambda$  - длина волны падающего излучения,  $S$  - область входной плоскости, в которой задано начальное поле.

В дальнейшем изложении учтём два обстоятельства: входная амплитуда задана не аналитическим выражением, а набором точек, получаемых при равномерной дискретизации по координатам с кусочно-постоянной аппроксимацией. Кроме того, временно не будем учитывать слагаемое  $i/(kR)$ , о котором будет отдельно сказано позже. Так как в пределах каждой ячейки разбиения входная амплитуда является постоянной, фактические вычисления проводятся по формуле:

$$U(x, y, z) = -\frac{ikz}{2\pi} \sum_{m,n} U_0[u_m, v_n] \times \int_{u_m}^{u_{m+1}} \int_{v_n}^{v_{n+1}} \frac{\exp\left[ik\sqrt{(u-x)^2 + (v-y)^2 + z^2}\right]}{(u-x)^2 + (v-y)^2 + z^2} dudv. \quad (2)$$

Далее сделаем ещё одно приближение – так как размер ячейки мал, а большого множителя в знаменателе нет, то будем считать, что знаменатель в пределах ячейки постоянный (возможность такого действия обсудим ниже). Поэтому формула (2) примет следующий вид:

$$U(x, y, z) = -\frac{ikz}{2\pi} \sum_{m,n} \frac{U_0[u_m, v_n]}{(\tilde{u}_m - x)^2 + (\tilde{v}_n - y)^2 + z^2} \int_{u_m}^{u_{m+1}} \int_{v_n}^{v_{n+1}} \exp\left[ik\sqrt{(u-x)^2 + (v-y)^2 + z^2}\right] dudv, \quad (3)$$

где  $(\tilde{u}_m, \tilde{v}_n)$  - центр ячейки.

В следующих пунктах мы изложим последовательные шаги увеличения скорости выполнения вычислений по формуле (3).

## 2. Прямой метод

В данном методе используется вычисление интегралов, входящих в формулу (3), непосредственно по квадратурным формулам (используется формула центральных прямоугольников). Отметим, что применено некоторое усовершенствование по сравнению с [5]. Оно состоит в том, что шаг интегрирования выбирается не на основе общей оценки для всей области, а выбирается для каждой ячейки, хотя, если мы находимся в условиях дальней зоны, выигрыш будет невелик. Примем в качестве шага интегрирования величину  $T/8$ ,  $T$  - период экспоненциальной подынтегральной функции. Значение знаменателя 8 выбрано как наименьшее число дроблений, при котором период описывается достаточно точно. Поэтому при применении данного метода (и других тоже) необходимо вычислить или хотя бы оценить период функции, которая не является точной синусоидой.

Будем рассуждать следующим образом. Для точной периодической функции  $\exp(ikx)$  частота равна  $k$  и период  $2\pi/k$ . Получить частоту можно также следующим образом. Имеем равенство

$$\frac{d}{dx} \exp(ikx) = ik \exp(ikx),$$

означающее, что частота – это коэффициент, на который умножается функция при дифференцировании, для точной периодической функции он является константой. Поступая по аналогии, получим равенство

$$\begin{aligned} \frac{d}{du} \exp\left(ik\sqrt{(u-x)^2 + (v-y)^2 + z^2}\right) = \\ = ik \frac{u-x}{\sqrt{(u-x)^2 + (v-y)^2 + z^2}} \times \\ \times \exp\left(ik\sqrt{(u-x)^2 + (v-y)^2 + z^2}\right). \end{aligned}$$

Из него следует, что величину

$$k \frac{|u-x|}{\sqrt{(u-x)^2 + (v-y)^2 + z^2}}$$

можно считать мгновенной частотой. Можно убедиться, что эта функция является возрастающей по отношению к величине  $u-x$ , поэтому частота будет максимальной, когда эта разность приобретает своё наибольшее значение в пределах ячейки. Поэтому имеем следующую формулу наименьшего периода:

$$T_{\min} = \frac{2\pi}{\omega_{\max}} = \frac{2\pi\sqrt{(u-x)_{\max}^2 + (v-y)^2 + z^2}}{k|u-x|_{\max}}. \quad (4)$$

На основе этой формулы и находится шаг интегрирования по переменной  $u$ , в силу симметрии функции аналогичное выражение будет при интегрировании по переменной  $v$ . В программной реали-

зации ячейки не разбиваются (отрезок интегрирования может содержать внутри себя границу ячеек), определяется только, какому значению входного поля соответствует центр ячейки. Поэтому, если шаг интегрирования окажется больше размера ячейки, вычисление производится корректно. Правда здесь может нарушиться условие правомочности перехода от (2) к (3), которое будет получено далее. Это явление можно было бы учесть в программе, но в этом нет необходимости по двум причинам – большой шаг может быть только в параксиальном случае, где нарушение условия маловероятно; быстродействие прямого метода очень мало (таблица 1), поэтому улучшать его не имеет смысла.

## 3. Переход к полярным координатам

Очевидный способ повышения быстродействия состоит в использовании того факта, что после замены переменных  $p = u-x$ ;  $q = v-y$  подынтегральная функция в (3) становится радиально симметричной. В этом случае целесообразно сделать переход к полярным координатам, что позволит аналитически взять интеграл по одной из переменных, что существенно повысит скорость вычислений, так как квадратурная формула будет применяться лишь для одномерного интеграла. Опишем это более подробно.

Нам необходимо вычислить интеграл следующего вида:

$$\begin{aligned} \int_{p_1}^{p_2} \int_{q_1}^{q_2} f(p^2 + q^2) dp dq, \\ p_1 = u_m - x \quad p_2 = u_{m+1} - x, \\ q_1 = v_n - y \quad q_2 = v_{n+1} - y, \end{aligned}$$

в котором область интегрирования – прямоугольник со сторонами, параллельными осям координат. В дальнейшем будем считать, что прямоугольник располагается в I четверти, так как к нему сводятся благодаря чётности подынтегральной функции остальные возможности расположения прямоугольника. После замены переменных получится интеграл  $\iint_D f(r) r dr d\phi$ , в котором область  $D$  – некоторый криволинейный четырёхугольник. Интеграл по переменной  $\phi$  в любом случае берётся аналитически, но, так как стороны четырёхугольника описываются разными уравнениями, исходный интеграл превращается в сумму интегралов по переменной  $r$ . Пределы интегрирования зависят от того, как расположен исходный прямоугольник. Есть четыре варианта:

- 1) прямоугольник не соприкасается с осями координат;
- 2) прямоугольник соприкасается с осью  $Oq$ ;
- 3) прямоугольник соприкасается с осью  $Op$ ;
- 4) прямоугольник соприкасается с обеими осями.

Варианты 1 и 4 также разделяются на два подварианта в зависимости от того, какая из вершин прямоугольника дальше от начала координат: левая верхняя или правая нижняя. Не будем описывать все

варианты, рассмотрим только вариант 1 для случая, когда правая нижняя вершина ближе левой верхней к началу координат.

При замене переменных  $p = r \cos \varphi$ ;  $q = r \sin \varphi$  прямоугольник ABCD преобразуется в криволинейный четырёхугольник A'B'C'D' (рис.1).

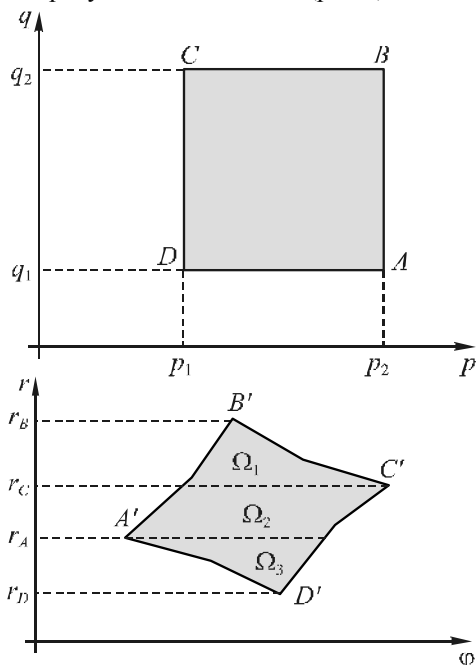


Рис.1. Преобразование области интегрирования при замене переменных

Преобразование границ описывается соотношениями:

$$AB : p = p_2 \rightarrow A'B' : r \cos \varphi = p_2$$

$$BC : q = q_2 \rightarrow B'C' : r \sin \varphi = q_2$$

$$CD : p = p_1 \rightarrow C'D' : r \cos \varphi = p_1$$

$$DA : q = q_1 \rightarrow D'A' : r \sin \varphi = q_1$$

Рисунок является схематичным: выпуклость или вогнутость границ не анализировалась.

Преобразование интеграла выражается формулами:

$$\int_{p_1}^{p_2} \int_{q_1}^{q_2} f(p^2 + q^2) dp dq = \iint_{\Omega} f(r) r dr d\varphi =$$

$$= \iint_{\Omega_1} (\cdot) + \iint_{\Omega_2} (\cdot) + \iint_{\Omega_3} (\cdot),$$

$$\iint_{\Omega_1} f(r) r dr d\varphi = \int_{r_C}^{r_B} f(r) r \left( \int_{\arccos(p_2/r)}^{\arcsin(q_2/r)} d\varphi \right) dr =$$

$$= \int_{r_C}^{r_B} f(r) r \left( \arcsin \frac{q_2}{r} - \arccos \frac{p_2}{r} \right) dr,$$

$$\iint_{\Omega_2} f(r) r dr d\varphi = \int_{r_A}^{r_C} f(r) r \left( \int_{\arccos(p_2/r)}^{\arccos(p_1/r)} d\varphi \right) dr =$$

$$= \int_{r_A}^{r_C} f(r) r \left( \arccos \frac{p_1}{r} - \arccos \frac{p_2}{r} \right) dr,$$

$$\iint_{\Omega_3} f(r) r dr d\varphi = \int_{r_D}^{r_A} f(r) r \left( \int_{\arcsin(q_1/r)}^{\arccos(p_1/r)} d\varphi \right) dr =$$

$$= \int_{r_D}^{r_A} f(r) r \left( \arccos \frac{p_1}{r} - \arcsin \frac{q_1}{r} \right) dr.$$

Если теперь использовать формулу  $\arccos x = \pi/2 - \arcsin x$ , раскрыть скобки и объединить интегралы с одинаковой подынтегральной функцией, то получим, что

$$\iint_{\Omega} f(r) r dr d\varphi =$$

$$= \int_{r_C}^{r_B} f(r) r \arcsin \frac{q_2}{r} dr + \int_{r_A}^{r_B} f(r) r \arcsin \frac{p_2}{r} dr -$$

$$- \int_{r_D}^{r_C} f(r) r \arcsin \frac{p_1}{r} dr - \int_{r_D}^{r_A} f(r) r \arcsin \frac{q_1}{r} dr -$$

$$- \frac{\pi}{2} \int_{r_C}^{r_B} f(r) r dr + \frac{\pi}{2} \int_{r_D}^{r_A} f(r) r dr.$$

Аналогично можно доказать, что и в остальных вариантах расположения исходной прямоугольной области интегрирования интеграл преобразуется в линейную комбинацию интегралов вида:

$$\int_{r_1}^{r_2} f(r) r dr \text{ и} \tag{5}$$

$$\int_{r_1}^{r_2} f(r) r \arcsin \frac{a}{r} dr. \tag{6}$$

В формулах (5) и (6)  $r_1, r_2$  - полярные радиусы некоторых вершин прямоугольника,  $a$  - одно из чисел  $p_1, p_2, q_1, q_2$ , причём гарантировано, что аргумент арксинуса не превосходит единицы.

Теперь в написанные выше выражения подставим функцию из (3)  $f(r) = \exp \left[ ik\sqrt{r^2 + z^2} \right]$  и получим, что результат является линейной комбинацией интегралов

$$\int_{r_1}^{r_2} \exp \left[ ik\sqrt{r^2 + z^2} \right] r dr \text{ и} \tag{7}$$

$$\int_{r_1}^{r_2} \exp \left[ ik\sqrt{r^2 + z^2} \right] r \arcsin \frac{a}{r} dr. \tag{8}$$

Интеграл в формуле (7) берётся в явном виде:

$$\int_{r_1}^{r_2} \exp \left[ ik\sqrt{r^2 + z^2} \right] r dr =$$

$$= \frac{1}{k^2} \left( \exp \left[ ik\sqrt{r^2 + z^2} \right] \cdot \left( 1 - ik\sqrt{r^2 + z^2} \right) \right) \Big|_{r_1}^{r_2}. \tag{9}$$

Интеграл (8) вычисляется по квадратурной формуле, при этом шаг выбирается равным 1/8 периода, так же как и в параграфе 1. Значение периода находится аналогично выводу формулы (4) и равно:

$$T_{\min} = \frac{2\pi}{\omega_{\max}} = \frac{2\pi\sqrt{r_{\max}^2 + z^2}}{kr_{\max}}. \tag{10}$$

Необходимо, однако, отметить, что между интегралами в (3) и (8) есть существенное отличие. В формуле (8) под знаком интеграла стоит не только экспоненциальная функция. В большинстве случаев это не оказывает влияния на выбор шага, так как шаг, вычисленный на основе периода экспоненты, обеспечивает нужную точность. Однако при малой частоте осцилляции  $\omega_{\max}$ , особенно когда прямоугольник расположен близко к началу координат, такой шаг получается слишком большим. Причина в том, что в этой ситуации множитель  $r \arcsin \frac{a}{r}$  изменяется на отрезке интегрирования быстрее экспоненты. Поэтому сделаем следующее преобразование:

$$\begin{aligned} \int_{r_1}^{r_2} \exp\left[ik\sqrt{r^2 + z^2}\right] r \arcsin \frac{a}{r} dr &\approx \\ &\approx \exp\left[ik\sqrt{r_s^2 + z^2}\right] \cdot \int_{r_1}^{r_2} r \arcsin \frac{a}{r} dr, \end{aligned} \tag{11}$$

где  $r_s = (r_1 + r_2) / 2$ .

Интеграл в правой части (11) берётся в явном виде:

$$\begin{aligned} \int r \arcsin \frac{a}{r} dr &= \left[ a / r = \sin x \right] = -a^2 \int \frac{x \cos x}{\sin^3 x} dx = \\ &= a^2 \frac{x}{2 \sin^2 x} + \frac{a^2}{2} \operatorname{ctgx} = a^2 \frac{\arcsin(a/r)}{2 \sin^2(\arcsin(a/r))} + \\ &+ \frac{a^2}{2} \operatorname{ctg}(\arcsin(a/r)) = \frac{r^2}{2} \arcsin \frac{a}{r} + \frac{a}{2} \sqrt{r^2 - a^2}. \end{aligned} \tag{12}$$

Итак, если шаг на основе периода экспоненты окажется велик, интеграл (8) вычисляем по следующей формуле:

$$\begin{aligned} \int_{r_1}^{r_2} \exp\left[ik\sqrt{r^2 + z^2}\right] r \arcsin \frac{a}{r} dr &\approx \\ &\approx \exp\left(ik\sqrt{r_s^2 + z^2}\right) \left( \frac{r^2}{2} \arcsin \frac{a}{r} + \frac{a}{2} \sqrt{r^2 - a^2} \right) \Big|_{r_1}^{r_2}. \end{aligned} \tag{13}$$

Теперь выясним, когда возможно применение преобразования (11). Можно доказать, используя выражение для погрешности квадратурных формул [6] и считая, что шаг интегрирования равен  $T/8$ ,

что, если отбросить общий множитель  $\frac{\pi^2}{16 \cdot 24}$ , отно-

сительная погрешность будет равна следующим значениям. Для интеграла от точной периодической функции  $e^{i\omega r}$  - числу 2 (точное значение), а для

функции  $r \arcsin \frac{a}{r}$  - величине  $\frac{0,679}{\frac{\pi}{2} r_1 \Delta \omega^2}$  (оценка),  $\Delta$  -

размер прямоугольника,  $\omega = 2\pi/T$ . Таким образом, если выполняется неравенство

$$2 < \frac{1,258}{\pi r_1 \Delta \omega^2}, \tag{14}$$

то преобразование (11) возможно. Отметим, что правая часть неравенства (14) оценена довольно грубо. Если  $r_1$  велико, то в числителе будет число большее, чем 1,258. Но в этом случае знаменатель тоже увеличится – см. формулу (10).

Применение перехода к полярным координатам существенно ускорило процесс вычислений: из таблицы 1 видно, что ускорение при малых значениях  $z$  может быть более 50 раз. Оно достигается за счёт того, что квадратурные формулы используются теперь для одномерного интеграла, несмотря на то, что алгоритм в целом стал более сложным.

#### 4. Использование асимптотической формулы в полярных координатах

При использовании алгоритма, описанного в пункте 3, наибольшие затраты времени происходят при вычислении на ячейках, далёких от начала координат. Это очевидно, так как шаг интегрирования на них меньше; в то же время их вклад в общую сумму сравнительно мал. Но из этого противоречия есть удачный выход – применение асимптотической формулы. Строго говоря, это название является условным, так как это не есть асимптотическое разложение в привычном его понимании. Мы учитываем то, что на удалённых ячейках арксинус мало изменяется:

$$\begin{aligned} \int_{r_1}^{r_2} \exp\left[ik\sqrt{r^2 + z^2}\right] r \arcsin \frac{a}{r} dr &\approx \\ &\approx \arcsin \frac{a}{r_s} \int_{r_1}^{r_2} \exp\left[ik\sqrt{r^2 + z^2}\right] r dr = \\ &= \frac{1}{k^2} \arcsin \frac{a}{r_s} \times \end{aligned} \tag{15}$$

$$\times \left[ \exp\left(ik\sqrt{r_2^2 + z^2}\right) \cdot \left(1 - ik\sqrt{r_2^2 + z^2}\right) - \exp\left(ik\sqrt{r_1^2 + z^2}\right) \cdot \left(1 - ik\sqrt{r_1^2 + z^2}\right) \right],$$

$$r_s = (r_1 + r_2) / 2.$$

Увеличение скорости очевидно – мы избавляемся от необходимости использовать квадратурные формулы. Необходимо только выяснить, когда такое вычисление не даёт слишком большой погрешности.

Для интеграла  $\int_{r_1}^{r_2} e^{i\omega r} dr$  относительная погреш-

ность вычисления по квадратурной формуле с шагом  $h = \frac{T}{8} = \frac{\pi}{4\omega}$  независимо от значений  $\omega, r_1, r_2$

равна  $\frac{\pi^2}{192}$ . Для интеграла от арксинуса при замене

подынтегральной функции константой  $\int_{r_1}^{r_2} \arcsin \frac{a}{r} dr \approx (r_2 - r_1) \arcsin \frac{a}{r_s}$  относительная погрешность, вычисленная по формуле (17), примерно равна  $\frac{a(r_2 - r_1)^2}{\arcsin \frac{a}{r_s}} \cdot \frac{2r_s^2 - a^2}{r_s^2(r_s^2 - a^2)^{3/2}}$ . Если она более чем в

два раза меньше погрешности при вычислении интеграла от экспоненты, то формула (15) применима. Таким образом, условие применимости формулы (15) выражается неравенством:

$$\frac{a(r_2 - r_1)^2}{\arcsin \frac{a}{r_s}} \cdot \frac{2r_s^2 - a^2}{r_s^2(r_s^2 - a^2)^{3/2}} < \frac{\pi^2}{384}. \tag{16}$$

Можно показать, что условие (16) выполняется, если ячейка отстоит от начала координат по обеим осям более чем на три своих размера.

Применение асимптотической формулы дало приемлемый результат в смысле точности, а скорость заметно возросла. По сравнению с алгоритмом из пункта 3 ускорение менее 2 раз в параксиальном случае и более 5 раз в непараксиальном. Однако имеется ещё один резерв повышения быстродействия, причём он не зависит от того, каким именно способом вычисляется интеграл на отдельной ячейке.

### 5. Запоминание интегралов по ячейкам в матрице с её последующим использованием

Ещё раз запишем сумму из формулы (3):

$$\sum_{m,n} \frac{U_0[u_m, v_n]}{(\tilde{u}_m - x)^2 + (\tilde{v}_n - y)^2 + z^2} \int_{p_m}^{p_{m+1}} \int_{q_n}^{q_{n+1}} \exp\left[ ik\sqrt{p^2 + q^2 + z^2} \right] dpdq, \\ p = u - x, \quad q = v - y.$$

Возможность повышения быстродействия обусловлена тем, что частичные интегралы в этой сумме повторяются многократно с одинаковыми значениями  $p_m, p_{m+1}, q_n, q_{n+1}$  для различных точек  $(x, y)$  выходной плоскости. Это означает, что если нужно вычислить одно значение  $U(x, y, z)$ , то факт повторения использовать невозможно. Но, если требуется найти целое поле  $U(x, y, z)$  и точки  $(x, y)$  расположены на дискретной сетке, то можно использовать тот факт, что величины  $p_m, q_n$  принимают конечное число значений. Это позволяет заранее вычислить для них интегралы, запомнить их в массиве и затем использовать их при вычислении вышеуказанной суммы.

Идея такого подхода взята из статьи [5], но в ней авторы не учли (во всяком случае, в статье об этом явно не сказано), что многократные повторения будут лишь в том случае, когда шаги дискретизации входного и выходного поля одинаковы. Если это ус-

ловие не выполнено, то повторений будет меньше и соответственно размер матрицы будет возрастать, что приведёт к уменьшению ускорения. Поэтому далее будем считать, что данное условие выполнено, тем более, что оно не является обременительным. (В принципе, можно использовать матрицу и в случае, когда шаги кратны друг другу, но при этом ускорение уменьшается и возрастает сложность программы, так как для разных соотношений шагов алгоритм несколько отличается.) Если входное поле имеет  $N_1 \times N_1$  точек, а выходное -  $N_2 \times N_2$  точек, то можно доказать, что размер матрицы будет  $(N_1 + N_2 - 1) \times (N_1 + N_2 - 1)$ .

Данный подход был программно реализован, было получено существенное ускорение, особенно после программной оптимизации цикла. До использования матрицы оптимизация не производилась, так как основное время тратилось на вычисление частичных интегралов. Было обнаружено, что почти всё время (даже после оптимизации) затрачивается на непосредственное вычисление выходного поля, а не на формирование матрицы. Однако это не значит, что преобразования, описанные в пунктах 3 и 4, оказываются бесполезными. Благодаря им получено ускорение большее, чем в [5], где частичные интегралы вычислялись по двумерным квадратурным формулам и получили упомянутое соотношение времён формирования матрицы и вычисления поля. С учётом этого соотношения, в отличие от [5], мы не стали использовать возможность вычислять не всю матрицу, а только её восьмую часть.

Строго говоря, даже при вычислении поля для одной точки  $(x, y)$  имеются повторяющиеся частичные интегралы, но при использовании матрицы это учитывать нецелесообразно. Дело в том, что будут большие программные издержки – повторяются не все частичные интегралы (кроме центральной точки  $U(0, 0, z)$ ).

### 6. Результаты вычислений

Здесь приводятся результаты вычислений, выполненных программой, реализующей описанные выше алгоритмы. Все эксперименты проводились при следующих значениях параметров: длина волны излучения 633 нм, размеры входной и выходной областей 4x4 мм с разбиением 51x51 точек. Конечно, такой шаг дискретизации в ряде случаев велик, а выходная область обычно берётся больше входной, но приводимая таблица предназначена только для сравнения времени выполнения различных алгоритмов. Входная функция – круглое отверстие. Расстояние между плоскостями является переменным. В таблице 1 приводится время выполнения программ и величина ускорения. Вычисления производились на компьютере, имеющем 512 МБайт оперативной памяти и тактовую частоту 2,01 ГГц.

Мы видим, что по сравнению с прямым методом алгоритм с использованием матрицы даёт ус-

корение как минимум в 34 раза, и более чем в 70000 раз в ближней зоне. В работе [5] получено для сопоставимых условий ускорение в 4 и более 1440 раз. У нас получено существенно большее ускорение благодаря использованию аналитических преобразований и, возможно, лучшей программной реализации. Также был повторён эксперимент, описанный в работе [3]. В этой статье сравнивался метод, предлагаемый авторами, и быстрые алгоритмы метода распространения плоских волн и аппроксимации Френеля. У нас было затрачено время 3,6 секунды, что существенно меньше полученного в [3] (5 минут для предлагаемого метода и 10 секунд для быстрых алгоритмов). Отметим, что эксперименты, описанные в [3], производились на более мощном компьютере: 4 Гбайт оперативной памяти и тактовая частота 2,8 ГГц.

Таблица 1. Время вычисления для области 4x4 мм с разбиением 51x51 точек

z, мм	Прямой метод (п.2)	Полярный 1 (п.3)	Полярный 2 (п.4)	Матрица (п.5)
2000	12 с	40с	36с	0,34с
1000	46 с	44с	36с	0,34с
500	3м03с	58с	35с	0,34с
250	12м10с	1м29с	35с	0,34с
125	≈50м	2м33с	36с	0,34с
70	≈400м	4м15с	37с	0,34с
30	≈1200м	9м23с	41с	0,34с
10	—	30м13с	55с	0,35с
1	—	156м	3м38с	0,43с

Точность работы программы оценивалась визуально следующим образом. Было проведено сравнение результатов, полученных матричным методом и прямым методом, реализованным в программе SimuLight.exe. Во всех случаях падающая волна была плоской. Примеры работы программы показаны на рисунках 2-4.

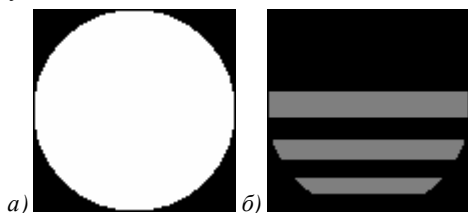


Рис.2. Входное поле:  $R=1$  мм, 100 отсчетов, а) интенсивность, б) фаза

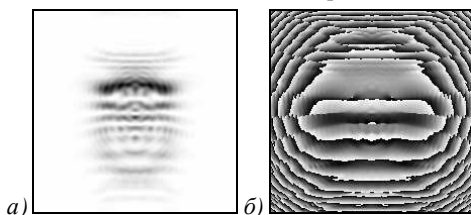


Рис.3. Выходное поле: расстояние  $z=300$  мм,  $R=2$  мм, 200 отсчетов, а) интенсивность (негатив), б) фаза. Использован матричный метод

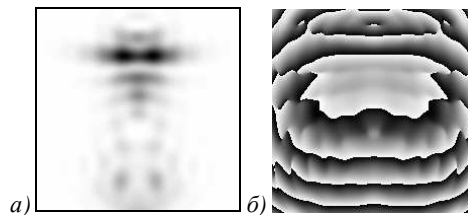


Рис.4. Выходное поле: расстояние  $z=1000$  мм,  $R=2$  мм, 200 отсчетов, а) интенсивность (негатив), б) фаза. Использован матричный метод

Обе программы дают результаты, примерно одинаковые визуально. Среднеквадратическое отклонение интенсивности равно 6,6% на расстоянии 300 мм и 4% на расстоянии 1000 мм. (Визуально такой же результат даёт программа Fresnel.exe, использующая приближение Френеля.) Другой способ сравнения следующий. Если расстояние между плоскостями очень мало, то выходное поле примерно равно входному. Такой пример показан на рисунках 5, 6. Наличие чёрной полосы на рисунке 6 а предсказывается теоретически.

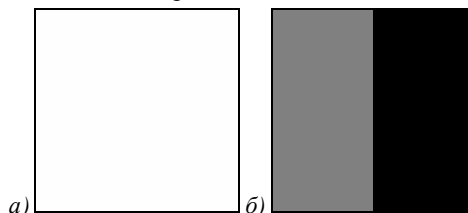


Рис.5. Входное поле: скачок фазы, 200 отсчетов, а) интенсивность, б) фаза

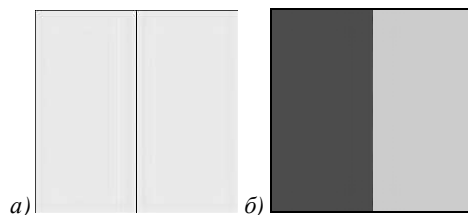


Рис.6. Выходное поле: расстояние  $z=0,01$  мм, 200 отсчетов, а) интенсивность, б) фаза

## 7. О возможности считать постоянной величину

$$1/R^2 + i/kR^3$$

Выше, в формуле (3) мы предположили, что величина  $1/((u-x)^2 + (v-y)^2 + z^2) + i/k((u-x)^2 + (v-y)^2 + z^2)^{3/2}$  в пределах

ячейки приблизительно постоянная, что позволяет заменить её значением в центре ячейки и вынести за знак интеграла. В данном пункте приводится условие, в каких случаях это возможно. При этом рассуждения с самого начала будут приводиться для случая, когда произведён переход к полярным координатам. Строго говоря, было получено и условие исходного перехода от (2) к (3), но мы не будем его приводить из-за того, что прямой метод слишком медленный, а само условие более жёсткое.

В самом начале было сказано, что слагаемое  $i/(kR)$  в подынтегральном выражении формулы (1) не учитывается. Условие, когда это можно делать,

достаточно очевидно -  $kR \gg 1$  [4]. Но, как мы увидим, это слагаемое целесообразно сохранять во всех случаях. Так сделано при программной реализации в методе с использованием матрицы (в остальных методах слагаемое  $i/(kR)$  не учитывается). Условия вынесения за знак интеграла рассмотрим вначале для каждого слагаемого в отдельности.

Пусть дана функция одной переменной  $f(r)$ . Если мы запишем следующее приближённое равенство:  $\int_{r_1}^{r_2} f(r)dr \approx f(r_s)(r_2 - r_1)$ ,  $r_s = (r_1 + r_2)/2$ , то его относительная погрешность примерно равна

$$\delta = \frac{f''(r_s)(r_2 - r_1)^2}{24f(r_s)}. \tag{17}$$

В нашем случае  $f(r) = \frac{1}{r^2 + z^2}$ , поэтому, исходя из (17), имеем:

$$\delta = \frac{(3r_s^2 - z^2)(r_2 - r_1)^2}{12(r_s^2 + z^2)^2}. \tag{18}$$

Разность  $r_2 - r_1 \leq \Delta$  при любом расположении ячейки (равенство может достигаться, если сторона ячейки лежит на координатной оси). Функция  $\frac{3r_s^2 - z^2}{(r_s^2 + z^2)^2}$ , рассматриваемая как функция одной переменной  $r_s$ , принимает наибольшее (по модулю) значение при  $r_s = 0$ . Таким образом,  $|\delta| \leq \frac{\Delta^2}{12z^2}$ .

Сравнивая это значение с величиной  $\frac{\pi^2}{384}$ , получаем, что вынесение функции  $\frac{1}{r^2 + z^2}$  за интеграл не приводит к ухудшению точности, если выполняется неравенство:

$$z > \frac{\sqrt{32}}{\pi} \Delta \approx 1,801\Delta. \tag{19}$$

Аналогично рассуждаем для второго слагаемого, в котором подынтегральная функция равна

$$f(r) = \frac{1}{(r^2 + z^2)^{3/2}}. \text{ Исходя из (17), имеем:}$$

$$\delta = \frac{(4r_s^2 - z^2)(r_2 - r_1)^2}{8(r_s^2 + z^2)^2}. \tag{20}$$

Анализируя правую часть (20), получим, что  $|\delta| \leq \frac{\Delta^2}{8z^2}$ , откуда получаем, что вынесение функции  $\frac{1}{(r^2 + z^2)^{3/2}}$  за интеграл не приводит к ухудшению точности, если выполняется неравенство:

$$z > \frac{\sqrt{48}}{\pi} \Delta \approx 2,205\Delta. \tag{21}$$

Теперь рассмотрим оба слагаемых вместе: вычислим погрешность для функции

$$\frac{1}{r^2 + z^2} + \frac{i}{k(r^2 + z^2)^{3/2}}. \tag{22}$$

Аналогично (18) и (20), имеем:

$$\delta = \frac{\left\{ 2 \frac{3r_s^2 - z^2}{(r_s^2 + z^2)^{3/2}} + \frac{3i}{k} \frac{4r_s^2 - z^2}{(r_s^2 + z^2)^2} \right\} (r_2 - r_1)^2}{24 \left\{ (r_s^2 + z^2)^{1/2} + i/k \right\}}. \tag{23}$$

Разность  $r_2 - r_1 \leq \Delta$ . Найдём максимум модуля оставшейся функции, обозначим её  $g(r_s)$ . Не будем выписывать его явный вид и приравнять к нулю производную. При рассмотрении слагаемых по отдельности было доказано, что значение  $g(r_s)$  при  $r_s = 0$  больше, чем при нуле производной. Поэтому предположим, что максимум модуля функции  $g(r_s)$  будет при  $r_s = 0$ .

$$\max |g(r_s)| = |g(0)| = \frac{1}{z^2} \sqrt{4 + \frac{5}{z^2 k^2 + 1}}.$$

Таким образом,

$$|\delta| \leq \frac{\Delta^2}{24z^2} \sqrt{4 + \frac{5}{z^2 k^2 + 1}}. \tag{24}$$

Отсюда видно, что погрешность лежит между значениями, соответствующими каждому слагаемому. При большом  $k$  получим  $|\delta| \leq \frac{\Delta^2}{12z^2}$ , а при малом  $kz$  будет  $|\delta| \leq \frac{\Delta^2}{8z^2}$ . В обоих случаях получим формулу для преобладающего слагаемого. Если нет явного преобладания одного из слагаемых, точное граничное значение для  $z$  получается из (24) решением уравнения 6 степени, которое сводится к кубическому. Можно также использовать оценку сверху – формулу (21). (Фактически в программе граница была увеличена до  $10\Delta$ .)

Если условие вынесения за знак интеграла не выполняется, то вместо (7) и (8) требуется вычислить два таких интеграла:

$$\int_{r_1}^{r_2} \exp[ik\sqrt{r^2 + z^2}] \left( \frac{1}{r^2 + z^2} + \frac{i}{k(r^2 + z^2)^{3/2}} \right) r dr \text{ и} \tag{25}$$

$$\int_{r_1}^{r_2} \exp[ik\sqrt{r^2 + z^2}] \left( \frac{1}{r^2 + z^2} + \frac{i}{k(r^2 + z^2)^{3/2}} \right) r \arcsin \frac{a}{r} dr. \tag{26}$$

Вычислим интеграл (25).

$$\begin{aligned}
& \int_{r_1}^{r_2} \exp \left[ ik\sqrt{r^2 + z^2} \right] \left( \frac{1}{r^2 + z^2} + \frac{i}{k(r^2 + z^2)^{3/2}} \right) r dr = \\
& = \left[ t = k\sqrt{r^2 + z^2} \right] = \int_{t_1}^{t_2} \frac{e^{it}}{t} dt + \frac{i}{k} \int_{t_1}^{t_2} \frac{e^{it}}{t^2} dt = \\
& = \int_{t_1}^{t_2} \frac{e^{it}}{t} dt + i \int_{t_1}^{t_2} \frac{e^{it}}{t^2} dt = \int_{t_1}^{t_2} \frac{e^{it}}{t} dt + \\
& + i \left( \frac{e^{it_1}}{t_1} - \frac{e^{it_2}}{t_2} + i \int_{t_1}^{t_2} \frac{e^{it}}{t} dt \right) = i \left( \frac{e^{it_1}}{t_1} - \frac{e^{it_2}}{t_2} \right),
\end{aligned} \quad (27)$$

где  $t_1 = k\sqrt{r_1^2 + z^2}$ ,  $t_2 = k\sqrt{r_2^2 + z^2}$ .

Таким образом, интеграл вычисляется в конечном виде. Именно поэтому второе слагаемое требуется сохранять во всех случаях.

Для интеграла (26), если выполняется условие (16), можно применить приём, аналогичный формуле (15):

$$\begin{aligned}
& \int_{r_1}^{r_2} \exp \left[ ik\sqrt{r^2 + z^2} \right] \left( \frac{1}{r^2 + z^2} + \frac{i}{k(r^2 + z^2)^{3/2}} \right) \times \\
& \times r dr \approx \arcsin \frac{a}{r} i \left( \frac{e^{it_1}}{t_1} - \frac{e^{it_2}}{t_2} \right) \arcsin \frac{a}{r_s},
\end{aligned} \quad (28)$$

### Заключение

В работе предложен эффективный способ вычисления интеграла Кирхгофа, который обеспечивает примерно одинаковую скорость вычислений независимо от расстояния между входной и выходной плоскостями. В ближней зоне, где невозможно использование приближения Френеля, выигрыш в скорости по сравнению с прямым вычислением по квадратурным формулам составляет тысячи раз. Аналогичный описанному подходу можно применить и в векторном случае. Так как подынтегральные функции в нём также зависят лишь от расстояния между точками входной и выходной плоскости, применение описанных преобразований возможно.

### Литература

1. **Totzeck, M.** Validity of the scalar Kirchhoff and Rayleigh-Sommerfeld diffraction theories in the near field of small phase objects // J. Opt. Soc. Am. – 1991.- Vol.8, N1. – P.27-32
2. **Melnikov, L.A.** The use of Kirchhoff approach for the calculation of the near field amplitudes of electromagnetic field / Melnikov L.A., Tsoy V.I. // Optics Communications. – 2005. – Vol.256 – P.1-9
3. **Veerman, J.A.C.**- Calculation of the Rayleigh-Sommerfeld diffraction integral by exact integration of the fast oscillating factor / J.A.C.Veerman, J.J.Rusch, H.P.Urbach // J. Opt. Soc. Am. – 2005.- Vol.22, N4. – P.636-646
4. **Виноградова, М.Б.** Теория волн / Виноградова М.Б., Руденко О.В., Сухоруков А.П.– М.: Наука, 1979. – 383 с.
5. **Балалаев, С.А.** Реализация быстрого алгоритма преобразования Кирхгофа на примере Бесселевых пучков / С.А.Балалаев, С.Н.Хонина. // Компьютерная оптика. – 2006. – Т.30 - С.69-73. – ISSN 0134-2452.
6. **Ланцош, К.** Практические методы прикладного анализа (глава VI) / К. Ланцош; пер. с англ. – М.: Государственное издательство физико-математической литературы, 1961. – 524 с. (Cornelius Lanczos. Applied Analysis, Prentice Hall, Inc., 1956).

### References

1. **Totzeck, M.** Validity of the scalar Kirchhoff and Rayleigh-Sommerfeld diffraction theories in the near field of small phase objects // J. Opt. Soc. Am. – 1991.- Vol.8, N1. – P.27-32
2. **Melnikov, L.A.** The use of Kirchhoff approach for the calculation of the near field amplitudes of electromagnetic field / Melnikov L.A., Tsoy V.I. // Optics Communications. – 2005. – Vol.256 – P.1-9
3. **Veerman, J.A.C.**- Calculation of the Rayleigh-Sommerfeld diffraction integral by exact integration of the fast oscillating factor / J.A.C.Veerman, J.J.Rusch, H.P.Urbach // J. Opt. Soc. Am. – 2005.- Vol.22, N4. – P.636-646
4. **Vinogradova, M.B.** Waves theory / Vinogradova M.B., Rudenko O.V., Suhorukov A.P.– Moscow: Science, 1979. – 383 p. – (in Russian).
5. **Balalayev, S.A.** Realization of fast algorithm of Kirchhoff's diffraction integral on an example of Bessel modes / S.A.Balalayev, S.N.Khonina. // Computer Optics. – 2006. – Vol.30 - P.69-73. – ISSN 0134-2452. – (in Russian)/
6. **Lanczos, C.** Applied Analysis / Prentice Hall, Inc., 1956. – 524 p.

## THE FAST WAY FOR CALCULATION OF FIRST CLASS RAYLEIGH-SOMMERFELD INTEGRAL

Andrey Vladimirovich Ustinov (leading programmer; e-mail: andr@smr.ru)

RAS Image Processing Systems Institute

### Abstract

Ways for acceleration of Kirchhoff integral owing to using of analytical transformations and the effective program realization are described in present article. Consecutive steps of the speed increase after every transformation are shown. Application conditions of different approximations are obtained. It should be noted that after an application of all transformations the time of program performance weakly depends on the distance between input and output planes.

**Key words:** Kirchhoff integral, paraxiality, quadrature formula, asymptotic expansion.

Поступила в редакцию 04.05.2009 г.