

The Use of Huygens' Equivalence Principle for Solving the Volume Integral Equation of Scattering

Weng Cho Chew, *Fellow, IEEE*, and Cai-Cheng Lu

Abstract—A new algorithm is introduced to solve the volume integral equation of scattering. A volume scatterer is first divided into N subscatterers. Then the subscatterers are divided into four groups, and the groups are in turn divided into four subgroups and so on. By using the idea found in many fast algorithms, a smaller problem can hence be nested within a larger problem. Moreover, by way of Huygens' equivalence principle, the scattering properties of a group of subscatterers in a volume can be replaced by a group of subscatterers distributed on a surface enclosing the volume. Based on this idea, we present an algorithm which solves the scattering problem by several stages, where at each stage the interaction matrix algorithm is first used to find the scattering solution of each subgroup of subscatterers. Subscatterers are then replaced by equivalent surface subscatterers which are used in the next stage. Consequently, this results in a reduction in the number of subscatterers at every stage. This algorithm can be shown to have a CPU time asymptotically proportional to $N^{1.5}$ for N subscatterers.

I. INTRODUCTION

FAST algorithms to solve integral equations have been studied by several investigators [1]–[5]. Many of these methods reduce the computational complexity of matrix-vector multiplications [1]–[3], [5] while others transform a dense matrix (e.g. found in the method of moments) to a sparse matrix [4]. Recently, we have developed a recursive aggregate T-matrix algorithm (RATMA) to solve volume integral equation with less computational complexity than a direct solution via a matrix equation with Gaussian elimination [6]–[13]. The method has a computational complexity of N^2 for two-dimensional scattering problems while it has a computational complexity of $N^{7/3}$ for three-dimensional scattering problems. Here N is the total number of unknowns needed to characterize the volume of the scatterer. These algorithms can also be applied to scattering by strips and patches [14]–[16].

In this paper, we introduce a new algorithm to solve volume integral equation with asymptotic computational complexity of $N^{1.5}$ in two dimensions and N^2 in three dimensions [17]. This is of less computational complexity than the recursive aggregate T-matrix algorithm (RATMA). This method makes use of Huygens' equivalence principle [18]–[20] in a nested manner. We shall call this the nested equivalence principle algorithm (NEPAL).

Manuscript received March 26, 1992; revised February 12, 1993. This work was supported by the Office of Naval Research under grant N00014-89-J1286 and by the Army Research Office under contract DAAL03-91-G-0339. The computer time was provided by the National Center for Supercomputing Applications (NCSA) at the University of Illinois, Urbana-Champaign.

The authors are with the Department of Electrical and Computer Engineering, University of Illinois, Urbana, IL 61801.

IEEE Log Number 9211246.

The nested equivalence principle algorithm is similar to many fast algorithms such as the FFT [21] and the nested dissection ordering approach in the finite element method [22]–[24]. The principal idea of this approach is to nest one algorithm within another so that a smaller problem is solved before a larger one. The size of the problem increases by a factor of 2^n at each stage, where n is the dimension of the problem. In this manner, an N unknown problem is solved in $\log_2 N$ steps. A nesting procedure of this type has also been used in [1], [2], and [5].

II. THE VOLUME INTEGRAL EQUATION

A volume integral equation has the form [18], [20]

$$\phi(\mathbf{r}) = \phi_{\text{inc}}(\mathbf{r}) + \underbrace{\int_V d\mathbf{r}' g_0(\mathbf{r} - \mathbf{r}') k_0^2 [\epsilon_r(\mathbf{r}') - 1] \phi(\mathbf{r}')}_{\phi_{\text{sca}}(\mathbf{r})} \quad (1a)$$

or

$$\phi(\mathbf{r}) = \phi_{\text{inc}}(\mathbf{r}) + \underbrace{\int_V d\mathbf{r}' \nabla g_0(\mathbf{r} - \mathbf{r}') [\epsilon_r^{-1}(\mathbf{r}') - 1] \cdot \nabla' \phi(\mathbf{r}')}_{\phi_{\text{sca}}(\mathbf{r})}, \quad (1b)$$

where $g_0(\mathbf{r} - \mathbf{r}')$ is $\frac{i}{4} H_0^{(1)}(k_0 |\mathbf{r} - \mathbf{r}'|)$ in two dimensions and $e^{ik_0 |\mathbf{r} - \mathbf{r}'|} / 4\pi |\mathbf{r} - \mathbf{r}'|$ in three dimensions. Equation (1a) is derivable from the differential equation

$$[\nabla^2 + k_0^2 \epsilon_r(\mathbf{r})] \phi(\mathbf{r}) = s(\mathbf{r}) \quad (2a)$$

while (1b) is derivable from

$$[\nabla \epsilon_r^{-1}(\mathbf{r}) \cdot \nabla + k_0^2] \phi(\mathbf{r}) = s(\mathbf{r}), \quad (2b)$$

assuming that the solutions satisfy radiation condition at infinity. In two dimensions, (2a) describes the scattering of E_z -polarized waves while (2b) is that of H_z -polarized waves.

One way to solve (1) is to discretize the scatterer into N_A discrete pieces and approximate the integrals in (1) with summations. Hence,

$$\phi_{\text{sca}}(\mathbf{r}) = \sum_{i=1}^{N_A} \Delta V w_i g_0(\mathbf{r} - \mathbf{r}'_i) k_0^2 [\epsilon_r(\mathbf{r}'_i) - 1] \phi(\mathbf{r}'_i) \quad (3a)$$

or

$$\phi_{\text{sca}}(\mathbf{r}) = \sum_{i=1}^{N_A} \Delta V w_i \nabla g_0(\mathbf{r} - \mathbf{r}'_i) [\epsilon_r^{-1}(\mathbf{r}'_i) - 1] \cdot \nabla' \phi(\mathbf{r}'_i). \quad (3b)$$

In the above w_i depends on the integration rule used [25].

In (3), the scattered field can be thought of as being scattered by N subscatterers, each with a scattering center at $\mathbf{r}'_i$. In (3a), the subscatterers are scattering like a monopole because they radiate via $g_0(\mathbf{r} - \mathbf{r}')$ while in (3b), the subscatterers are scattering like a dipole because they radiate via $\nabla g_0(\mathbf{r} - \mathbf{r}'_i)$. Obviously, equation (3) will be a good approximation of the scattered field if the discretization density is high enough.

The general expression for the scattered field is in terms of a linear superposition of monopole and dipole fields, each centered at $\mathbf{r}'_i$. Moreover, these monopole and dipole fields are expressible in terms of cylindrical harmonics in two dimensions and spherical harmonics in three dimensions. Hence, in general, the scattered field can be expressed as [12], [20, p. 464]

$$\phi_{\text{sca}}(\mathbf{r}) = \sum_{i=1}^{N_A} \psi^t(k_0, \mathbf{r}_i) \cdot \mathbf{b}_i, \quad (4)$$

where $\mathbf{r}_i = \mathbf{r} - \mathbf{r}'_i$, and $\psi^t(k_0, \mathbf{r}_i)$ is a row vector containing cylindrical harmonics for two dimensions, and spherical harmonics in three dimensions.¹ Its element $\psi_m(\mathbf{r}_i)$ is also defined in the Appendix. The column vector $\mathbf{b}_i$ contains the magnitudes of the harmonic expansions. When the size of the subscatterer is small, say about 0.1λ in diameter, the harmonic expansions need only contain the monopole and the dipole terms. For monopole subscatterers, $\mathbf{b}_i$ is a one-component vector corresponding to the monopole harmonics, and for dipole subscatterers, $\mathbf{b}_i$ is a two-component vector in two dimensions corresponding to two harmonics ($e^{i\phi}$ and $e^{-i\phi}$), and a three-component vector in three dimensions. The unknown $\mathbf{b}_i$ can be found by imposing the boundary conditions on all the N subscatterers.

III. HUYGENS' EQUIVALENCE PRINCIPLE

Huygens' equivalence principle states that the scattered field from a scatterer can be replaced by the field radiated by equivalent sources on a surface S which encloses the scatterer. Stated mathematically, this is [20, p. 29]

$$\phi_{\text{sca}}(\mathbf{r}) = - \oint_S dS' \hat{n}' \cdot \underbrace{[\phi_{\text{sca}}(\mathbf{r}') \nabla' g_0(\mathbf{r} - \mathbf{r}')]_{\text{dipole}}}_{\text{dipole}} - \underbrace{g_0(\mathbf{r} - \mathbf{r}') \nabla' \phi_{\text{sca}}(\mathbf{r}')}_{\text{monopole}}. \quad (5)$$

The above is also known as Green's theorem [18]. Since $\hat{n}' \cdot \nabla' g_0(\mathbf{r} - \mathbf{r}')$ is a field due to an $\hat{n}'$ -directed dipole at $\mathbf{r}'$ while $g_0(\mathbf{r} - \mathbf{r}')$ is a field due to a monopole (or point charge) at $\mathbf{r}'$, the first term in (5) can be attributed to a surface dipole-layer source while the second term in (5) can be related to a surface charge source.

¹Throughout this paper, we will use the notation $\mathbf{r}_i = \mathbf{r} - \mathbf{r}'_i$, and the "prime" will be used to denote a source location.

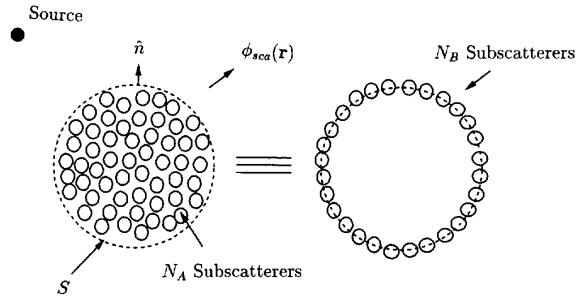


Fig. 1. Huygens' equivalence principle is used to replace the volume scattering centers by surface scattering centers.

In an actual numerical implementation, the integral in (5) is discretized and becomes

$$\phi_{\text{sca}}(\mathbf{r}) = - \sum_{m=1}^{N_B} \Delta S w_m [\phi_{\text{sca}}(\mathbf{r}'_m) \hat{n}' \cdot \nabla' g_0(\mathbf{r} - \mathbf{r}'_m) - \hat{n}' \cdot \nabla' \phi_{\text{sca}}(\mathbf{r}'_m) g_0(\mathbf{r} - \mathbf{r}'_m)], \quad (6)$$

where w_m depends on the integration rule used. In this numerical approximation, the scattered field from N_A subscatterers is replaced by that from N_B subscatterers on the surface S (Fig. 1).

Since the scattering centers in (6) are monopole and dipole centers whose scattered fields are expressible in terms of cylindrical or spherical harmonics, (6) can be written in general as

$$\phi_{\text{sca}}(\mathbf{r}) = \sum_{m=1}^{N_B} \psi^t(k_0, \mathbf{r}_m) \cdot \mathbf{c}_m, \quad (7)$$

where $\mathbf{r}_m = \mathbf{r} - \mathbf{r}'_m$, and $\mathbf{c}_m$ depends on $\phi_{\text{sca}}(\mathbf{r}'_m)$ and $\hat{n}' \cdot \nabla' \phi_{\text{sca}}(\mathbf{r}'_m)$. In general, when the subscatterers are small enough, at most three harmonics are needed in $\psi^t(k_0, \mathbf{r}_m)$ to account for the monopole and dipole scattered fields from each scattering center. Hence, the unknown $\mathbf{c}_m$ is also a three-component vector. The above procedure has the important effect of reducing the number of scattering centers of a scatterer by replacing the volume subscatterers by surface subscatterers. This is one of the key points for the success of NEPAL.

Since $\phi_{\text{sca}}(\mathbf{r}'_m)$ is a linear functional of $\mathbf{b}_i$ in (4), a general linear relationship can be established between $\mathbf{c}_m$ in (7) and $\mathbf{b}_i$ in (4). It can be shown that

$$\mathbf{c}_m = \sum_{i=1}^{N_A} \bar{\mathbf{h}}_{mi}^{(o)} \cdot \mathbf{b}_i \quad (8a)$$

or

$$\phi_{\text{sca}}(\mathbf{r}) = \sum_{m=1}^{N_B} \psi^t(k_0, \mathbf{r}_m) \cdot \sum_{i=1}^{N_A} \bar{\mathbf{h}}_{mi}^{(o)} \cdot \mathbf{b}_i. \quad (8b)$$

Since $\mathbf{c}_m$ and $\mathbf{b}_i$ are three-component vectors, $\bar{\mathbf{h}}_{mi}^{(o)}$ is a 3×3 matrix which is dependent on the location of the subscatterers. The explicit form for $\bar{\mathbf{h}}_{mi}^{(o)}$ in two dimensions is given in the Appendix.

The above represents the use of Huygens' equivalence principle to alter the scattering centers of the scattered waves. A similar idea could be applied to the incident wave so that the scatterer, instead of being illuminated by a source far from the scatterer, is illuminated by some equivalent sources much closer to the scatterer.

To this end, we express the incident field impinging on the N_A subscatterers in general as one generated by a source (monopole or dipole, or combination thereof) located at $\mathbf{r}'_s$. Hence,

$$\phi_{\text{inc}}(\mathbf{r}) = \psi^t(k_0, \mathbf{r}_s) \cdot \mathbf{a}_s, \quad (9)$$

where $\mathbf{r}_s = \mathbf{r} - \mathbf{r}'_s$ and $\mathbf{r}'_s$ is a position vector denoting the location of the source in Fig. 1. Equation (9) is valid everywhere outside the source region. Using the addition theorem, this incident field can be expressed as being impinging on the j th subscatterers in the N_A subscatterers, or [12], [20, p. 464]

$$\phi_{\text{inc}}(\mathbf{r}) = \mathcal{R}g\psi^t(k_0, \mathbf{r}_j) \cdot \bar{\alpha}_{js} \cdot \mathbf{a}_s. \quad (10)$$

Here, $\mathcal{R}g\psi^t(k_0, \mathbf{r}_j)$ is a vector containing the regular part of the harmonics involving Bessel functions and $\mathbf{r}_j = \mathbf{r} - \mathbf{r}'_j$. $\bar{\alpha}_{js}$ is a translation matrix [12], [20, p. 464] that allows the change of coordinates of the wave functions (see also the Appendix). Since the expression (10) need only be valid over the smallest circular region containing the j th subscatterer, which is small compared with the wavelength, the number of harmonics needed in expanding (10) is at most three. Hence, $\bar{\alpha}_{js}$ is at most a 3×3 matrix.

Similarly, by using Huygens' equivalence principle, the incident wave, instead of coming from one source as in (9), now comes from N_B equivalent sources. Hence,

$$\phi_{\text{inc}}(\mathbf{r}) = \mathcal{R}g\psi^t(k_0, \mathbf{r}_j) \cdot \sum_{m=1}^{N_B} \bar{h}_{jm}^{(i)} \cdot \bar{\alpha}_{ms} \cdot \mathbf{a}_s. \quad (11)$$

When these equivalent sources are small compared with the wavelength, the number of harmonics needed in (11) is at most three, similar to the case in (7). By similar argument given to $\bar{\alpha}_{js}$ in (10), $\bar{\alpha}_{ms}$ in (11) is a 3×3 matrix, and hence, $\bar{h}_{jm}^{(i)}$ is a 3×3 matrix. It is explicitly given in the Appendix.

Equation (11) can be used to shift the centers of the incident field. Note that (8b) is valid everywhere outside S and (11) is valid everywhere inside S within numerical approximations. These two procedures of shifting the scattering centers and the centers of the incident field are important for the subsequent implementation of NEPAL.

IV. REDUCTION OF INTERACTION MATRICES

If the scattering solution is known, then the scattered field amplitudes $\mathbf{b}_i$ in (4) are known. Moreover, they are linearly related to the incident field via interaction matrices [26]. Hence, we can rewrite (4) as

$$\phi_{\text{sca}}(\mathbf{r}) = \sum_{i=1}^{N_A} \psi^t(k_0, \mathbf{r}_i) \cdot \sum_{j=1}^{N_A} \bar{I}_{ij(N_A)} \cdot \bar{\alpha}_{js} \cdot \mathbf{a}_s, \quad (12)$$

where $\mathbf{r}_i = \mathbf{r} - \mathbf{r}'_i$, $\bar{I}_{ij(N_A)}$ is the interaction matrix that accounts for wave interaction between the i th and j th subscatterers, $\mathbf{a}_s$ denotes the amplitudes of the incident waves in (9), and $\bar{\alpha}_{js}$ is the translation matrix that changes the coordinates of the wave from the s coordinates to the j coordinates (see the Appendix). Note that there are N_A^2 interaction matrices in (12).

Using (8b) and (11), which have been used to rewrite (7) and (10), we can rewrite the above as

$$\phi_{\text{sca}}(\mathbf{r}) = \sum_{m=1}^{N_B} \psi^t(k_0, \mathbf{r}_m) \cdot \sum_{i=1}^{N_A} \bar{h}_{mi}^{(o)} \cdot \sum_{j=1}^{N_A} \bar{I}_{ij(N_A)} \cdot \sum_{m'=1}^{N_B} \bar{h}_{jm'}^{(i)} \cdot \bar{\alpha}_{m's} \cdot \mathbf{a}_s, \quad (13)$$

where $\mathbf{r}_m = \mathbf{r} - \mathbf{r}'_m$. The above can be rewritten as

$$\phi_{\text{sca}}(\mathbf{r}) = \sum_{m=1}^{N_B} \psi^t(k_0, \mathbf{r}_m) \cdot \sum_{m'=1}^{N_B} \bar{I}_{mm'(N_B)} \cdot \bar{\alpha}_{m's} \cdot \mathbf{a}_s, \quad (14)$$

where

$$\bar{I}_{mm'(N_B)} = \sum_{i=1}^{N_A} \bar{h}_{mi}^{(o)} \cdot \sum_{j=1}^{N_A} \bar{I}_{ij(N_A)} \cdot \bar{h}_{jm'}^{(i)}. \quad (15)$$

In (14), the number of interaction matrices is reduced from N_A^2 to N_B^2 , where $N_B < N_A$. Hence, by replacing volume scatterers with surface scatterers, the number of scattering centers used to describe the scattering process is greatly reduced. Furthermore, there is no violation of the addition theorem in using this equivalence principle.

V. THE NESTED EQUIVALENCE PRINCIPLE ALGORITHM (NEPAL)

We shall describe the nested equivalence principle algorithm (NEPAL) verbally first in this section, followed by a mathematical description in the next section.

The nested equivalence principle algorithm is similar in spirit to the fast Fourier transform and nested dissection ordering in the finite element method [21]–[24]. The idea is to divide the original problem into subgroups. The subgroups are then further divided into sub-subgroups until the lowest level of subgroups is reached. In two dimensions, the lowest level of subgroups is in groups of four while in three dimensions, it is in groups of eight. (For 1-D FFT, it is in groups of two.) The goal of NEPAL is to find a scattering solution that accounts for all the interactions between the subscatterers. To simplify the discussion, the following is for two dimensions using a rectilinear grid, even though a rectilinear grid is not mandatory. Similar discussion can be extended to three dimensions.

It is assumed that the scattering solution of each subscatterer is known. The solution could be found for instance by the T-matrix method or its like [12]. Next, the scattering problem involving the lowest level of subgroups of subscatterers is solved. In two dimensions, even though a subgroup in the lowest level consists of four subscatterers, we shall not start at the group of four because Huygens' equivalence principle, to

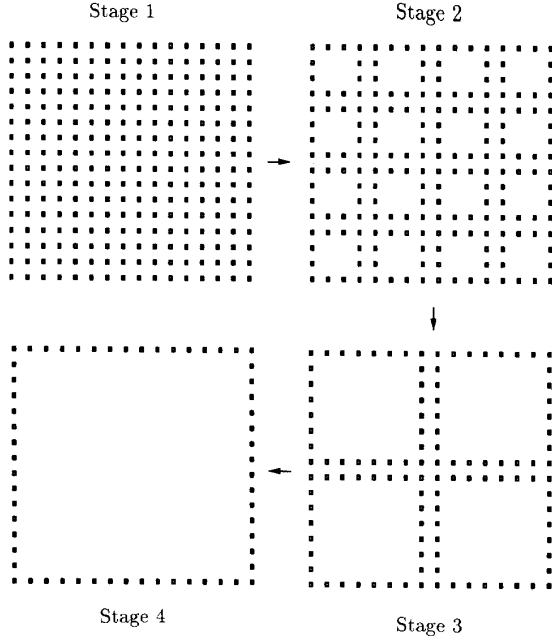


Fig. 2. By nesting a smaller problem within a larger one, Huygens' equivalence principle is used to reduce gradually the number of scattering centers by surface scattering centers. Scattering problems are also solved at each stage.

be discussed next, cannot be used here. Instead, we start with the group of 16 subscatterers in our first stage.

Hence, in the first stage, the scattering solutions for groups of 16 subscatterers are obtained. These scattering solutions would account for only interactions between the 16 subscatterers in the group but would not account for interactions of the subscatterers within the group with any other subscatterers outside the group. Hence, they do not represent the correct solution yet. After the scattering solution from a group of 16 subscatterers is known, Huygens' equivalence principle is used to replace the 16 scattering centers with 12 scattering centers. The procedure is repeated $N/16$ times, representing the number of subgroups in the first stage. This is illustrated by stage 2 in Fig. 2. Note that by so doing, the number of scattering centers is reduced at the first stage.

In the next stage, the scattering solutions are solved for groups of four subgroups, each consisting of 12 subscatterers, implying a total of 48 subscatterers. Again, after the solution is sought, Huygens' equivalence principle is used to reduce the number of scattering centers from 48 to 28. This procedure is repeated $N/(16 \cdot 4)$ times. This is illustrated by stage 3 in Fig. 2. The process is repeated until the scattering solution for the entire scatterer is obtained whereby interactions between all the subscatterers are accounted for.

As the subgroup is growing from stage to stage, the number of elements in each subgroup is growing as the sequence $4 \cdot 4$, $4 \cdot 12$, $4 \cdot 28$, $4 \cdot 60$, $4 \cdot 124, \dots$. The important thing to note here is that at each stage the number of groups is reduced by a factor of 4, while the number of unknowns is roughly increased by a factor of 2, as denoted by the above sequence. Moreover, the problem is solved in $\log_4(N/4)$ stages.

In the nested equivalence principle algorithm, a scattering solution for four groups of scatterers is required when the scattering solution of each group is known. This calls for a recursive algorithm which provides the solution of the larger groups from the smaller groups. To this end, we shall use the interaction matrix approach (IMA) [26]. It is to be noted that the interaction matrix approach may not be the only approach possible for this implementation.

VI. THE INTERACTION MATRIX APPROACH (IMA)

Notice that in the above algorithm, it is necessary to derive the scattering solution of four groups of subscatterers if the solution of each of the groups is known. We shall discuss finding the scattering solution of four groups of subscatterers assuming the solutions of each group of subscatterers are known (see Fig. 3).

First, assume that the number of subscatterers of each group equals N_{Bi} , where $i = 1, \dots, 4$, and that the scattering solution of each group is known.² Given the scattering solution of a group of N_{Bi} subscatterers, the scattering solution can be written using the interaction matrix approach (IMA) as [26]

$$\phi_{sca}^i(\mathbf{r}) = \sum_{l=1}^{N_{Bi}} \psi^l(k_0, \mathbf{r}_l) \cdot \sum_{m=1}^{N_{Bi}} \bar{I}_{lm(N_{Bi})} \cdot \bar{\alpha}_{ms} \cdot \mathbf{a}_s \quad (16)$$

With four groups of (16), we are to obtain the scattering solution for a larger group, N_c , of subscatterers, where $N_c = \sum_{i=1}^4 N_{Bi}$. Hence, we seek the solution

$$\phi_{sca}(\mathbf{r}) = \sum_{l=1}^{N_c} \psi^l(k_0, \mathbf{r}_l) \cdot \sum_{m=1}^{N_c} \bar{I}_{lm(N_c)} \cdot \bar{\alpha}_{ms} \cdot \mathbf{a}_s \quad (17)$$

This solution now accounts for the interactions between all the subscatterers in the four groups.

The above problem can be solved by the matrix-partitioning method. Alternatively, we can solve it with the interaction matrix approach (IMA) [26]. To this end, it is more suggestive to write (16) as a one-scatterer solution and (17) as a four-scatterer solution. Hence, (16) can be expressed as

$$\phi_{sca}^i(\mathbf{r}) = \Psi^t(k_0, \mathbf{r}_i) \cdot \bar{T}_{i(1)} \cdot \bar{A}_{is} \cdot \mathbf{a}_s, \quad i = 1, \dots, 4, \quad (18)$$

where

$$[\Psi^t(k_0, \mathbf{r}_i)]_l = \psi^l(k_0, \mathbf{r}_l), \quad l = 1, \dots, N_{Bi} \quad (19a)$$

$$[\bar{T}_{i(1)}]_{lm} = \bar{I}_{lm(N_{Bi})}, \quad l, m = 1, \dots, N_{Bi} \quad (19b)$$

$$[\bar{A}_{is}]_l = \bar{\alpha}_{ls}, \quad l = 1, \dots, N_{Bi}. \quad (19c)$$

Hence, when four groups of scatterers are present, they resemble a four-scatterer problem, and (17) could be expressed as

$$\phi_{sca}(\mathbf{r}) = \sum_{i=1}^4 \Psi^t(k_0, \mathbf{r}_i) \cdot \sum_{j=1}^4 \bar{T}_{ij(4)} \cdot \bar{A}_{js} \cdot \mathbf{a}_s, \quad (20)$$

²Even though in this work, N_{Bi} is the same for different i 's, they need not be the same.

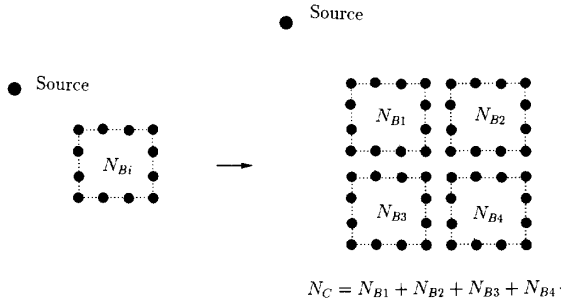


Fig. 3. The interaction matrix approach is used to account for the interactions between four groups of subscatterers given that the interactions for subscatterers within each of the groups are known.

where

$$\begin{aligned} [\bar{T}_{ij(4)}]_{pq} &= \bar{I}_{(i-1)N_{B,i-1}+p,(j-1)N_{B,j-1}+q(N_C)} \\ p &= 1, \dots, N_{Bi}, \quad q = 1, \dots, N_{Bj}, \\ i, j &= 1, \dots, 4. \end{aligned} \quad (21)$$

Therefore, once $\bar{T}_{ij(4)}$ is known, $\bar{I}_{lm(N_C)}$ is known. $\bar{T}_{ij(4)}$ can be obtained from $\bar{T}_{i(1)}$ using a recursive algorithm via the interaction matrices approach (IMA) [26]. Reproduced below, it is

$$\begin{aligned} \bar{T}_{n+1,j(n+1)} &= \bar{D}^{-1} \cdot \bar{T}_{n+1(1)}, \\ \bar{D} &= \left[\bar{I} - \bar{T}_{n+1(1)} \cdot \sum_{i=1}^n \bar{A}_{n+1,i} \cdot \sum_{j=1}^n \bar{T}_{ij(n)} \cdot \bar{A}_{j,n+1} \right] \end{aligned} \quad (22a)$$

$$\bar{T}_{n+1,j(n+1)} = \bar{T}_{n+1,n+1(n+1)} \cdot \left[\sum_{i=1}^n \bar{A}_{n+1,i} \cdot \bar{T}_{ij(n)} \right] \quad (22b)$$

$$\bar{T}_{i,n+1(n+1)} = \left[\sum_{j=1}^n \bar{T}_{ij(n)} \cdot \bar{A}_{j,n+1} \right] \cdot \bar{T}_{n+1,n+1(n+1)} \quad (22c)$$

$$\bar{T}_{ij(n+1)} = \bar{T}_{ij(n)} + \left[\sum_{j'=1}^n \bar{T}_{ij'(n)} \cdot \bar{A}_{j',n+1} \right] \cdot \bar{T}_{n+1,j(n+1)}, \quad (22d)$$

where $i, j = 1, \dots, n$, $n = 1, 2, 3$, and

$$\begin{aligned} [\bar{A}_{j,n+1}]_{lm} &= \bar{\alpha}_{lm}, \quad l = 1, \dots, N_{Bj}, \\ m &= 1, \dots, N_{B,n+1}. \end{aligned} \quad (23)$$

IMA has a computational complexity of $O(N^3)$ for solving an N -unknown problem.

VII. COMPUTATIONAL COMPLEXITY

For simplicity, rectilinear grid is used for the following discussion. It is observed that for this grid, the use of Huygens' equivalence principle reduces the number of scattering center from 16 to 12 in each group, before grouping by four is done. From the discussion in Section V, the number of scattering centers in a group increases as the sequence $4 \cdot 4, 4 \cdot 12, 4 \cdot 28, 4 \cdot 60, 4 \cdot 124, \dots$. This sequence is described by $4n_i$,

where n_i satisfies the difference equation $n_i = 2n_{i-1} + 4$ with initial value $n_1 = 4$. The solution for n_i is shown easily to be $n_i = 4(2^i - 1)$.

We assume N scattering centers in the original problem, each of which is at most a dipole scatterer, so that we can approximate the scattered field from each center by M harmonics, where M is at most 3. Hence, there are all together NM unknowns. Furthermore, we assume that N is a power of 4 so that at the lowest level there are $N/16$ subgroups, each with 16 scattering centers. With $16M$ unknowns for each group, the number of operations needed to solve for them is $(16M)^3$ with an $O(n^3)$ algorithm repeated $N/16$ times. Hence, the total operation count at the lowest level is proportional to $(N/16)(16M)^3$.

Consequently, at the i th stage, the number of unknowns is $4n_i M$ per subgroup with $N/(4 \cdot 4^i)$ subgroups. Hence, the operation count needed at the i th stage is $N/(4 \cdot 4^i)(4n_i M)^3$. There are a total of $\log_4(N/4)$ stages in this method. Consequently, the CPU time is proportional to

$$T \sim 16NM^3 \sum_{i=1}^{\log_4(N/4)} \frac{n_i^3}{4^i} = 4^5 NM^3 \sum_{i=1}^{\log_4(N/4)} \frac{(2^i - 1)^3}{4^i}. \quad (24)$$

Hence, keeping the leading order term only, we have

$$T \sim 2 \cdot 4^5 NM^3 \sum_{n=0}^{\log_4(N/4)-1} 2^n = 2 \cdot 4^5 NM^3 \frac{1 - 2^{\log_4(N/4)}}{1 - 2}. \quad (25)$$

When $N \rightarrow \infty$, it is seen that

$$T \sim 1024N^{1.5}M^3, \quad N \rightarrow \infty. \quad (26)$$

Therefore, this is an asymptotically an $N^{1.5}$ algorithm.³ It is seen that the multiplicative constant in (26) is quite large. However, due to the systematic nature of this algorithm, a vector machine or a massively parallel processors can be used to reduce the constant in (26). The analysis of the memory requirement shows that it requires $O(N \log N)$ memory storage.

VIII. NUMERICAL RESULTS

Using this algorithm, we have developed a program to compute the scattering solution of square dielectric cylinders for both E_z and H_z polarized incident waves. The results agree well with that of the method of moments (MOM) [28], as shown in Figs. 4 and 5.

As for the computation complexity, Fig. 6 shows the comparison of CPU time (on CRAY-2) of this method (NEPAL) with the recursive aggregated T-matrix algorithm (RATMA) and the method of moments (MOM). Since RATMA is highly efficient, NEPAL is only slightly better than RATMA within the range of unknowns we have shown. However, the curve indicates that NEPAL has an asymptotic complexity of $N^{1.5}$, albeit with a large proportionality constant as shown by (26). RATMA has an asymptotic complexity of N^2 , as shown in [12].

³ A similar analysis shows that it is N^2 in three dimensions.

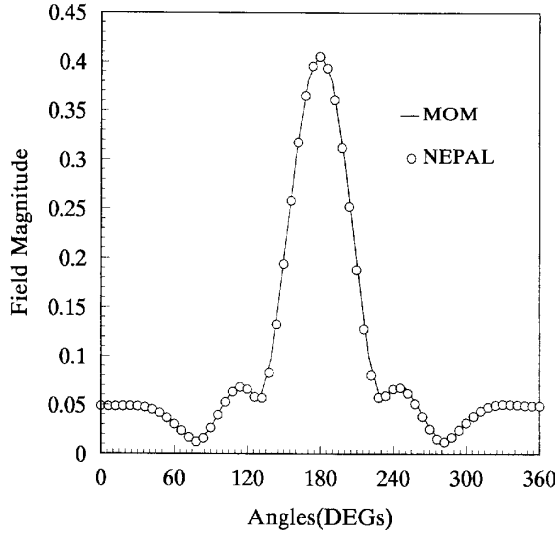


Fig. 4. Scattered field of a square dielectric cylinder for E_z polarized plane wave incidence at $\theta = 0^\circ$. Side length $= 1\lambda$, $\epsilon_r = 2.0$, $f = 300$ (MHz). Field points are on the circle of $r = 30\lambda$. The cylinder is divided into 64 small subscatterers.

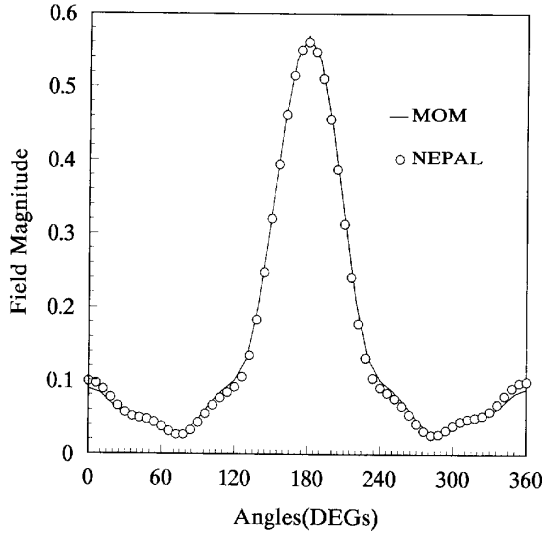


Fig. 5. Scattered field of a square dielectric cylinder for H_z polarized plane wave incidence at $\theta = 0^\circ$. Side length $= 0.9\lambda$, $\epsilon_r = 2.0$, $f = 300$ (MHz). Field points are on the circle of $r = 10\lambda$. The cylinder is divided into 64 small subscatterers.

For small N , MOM spends most of its time in calculating the matrix elements with a complexity of N^2 . For large N , the dominant time in the MOM is the matrix inversion time. We show on the same plot various methods of factorizing a complex matrix. CMINV is a program we have modified from SSPLIB for complex matrices, GAUSSJ is a program we have modified from [29] for complex matrices, while for NAGLIB, we use special CRAY purpose programs F01NAF and F04NAF for factorization and back-substitution written

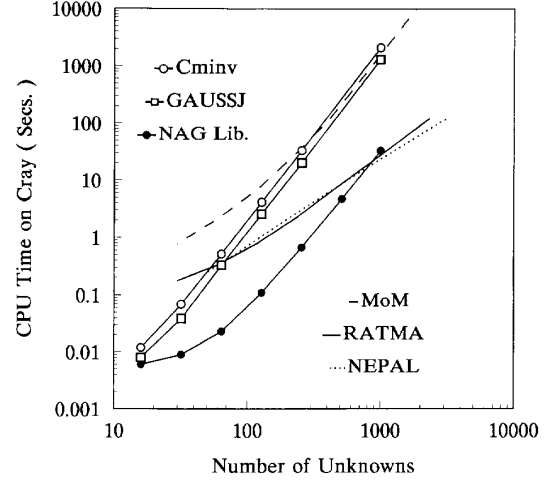


Fig. 6. Comparison of CPU time of NEPAL with RATMA and MOM, and various methods of matrix inversion.

in assembly language specially tailored for the CRAY-2. The use of NAGLIB can reduce the matrix factorization time considerably. However, it does not reduce the complexity of matrix factorization, as can be seen from the curve, which asymptotically has a slope of 3. It is possible that by writing special-purpose programs for NEPAL and RATMA for the CRAY, their constants of proportionality in their CPU time can be reduced, but not their complexity.

Notice that NEPAL and RATMA both outperform GAUSSJ and CMINV (which are portable codes) in solving the scattering problem. This is true even for a small number of unknowns. NEPAL is much faster than MOM and it is at least as efficient as RATMA. For small N , its performance is not superior to RATMA, even though in theory NEPAL is of reduced computational complexity compared to RATMA. The reason is as follows: in NEPAL, we attempt to reduce the number of scattering centers by a factor of 2 using Huygens' equivalence principle. For a small number of unknowns, the reduction of the number of subscatterers at each stage is not exactly a factor of 2. The reduction ratio at the first stage is $R_1 = 16/12 = 1.33$. The reduction ratios are subsequently $R_2 = 1.714$, $R_3 = 1.867$, $R_4 = 1.936, \dots$, which quickly approach 2. However, it is expected that when the number of unknowns becomes very large, this algorithm will demonstrate more of its advantage.

IX. CONCLUSIONS

We have presented in this paper an efficient algorithm for computing the scattering solution of large objects. The algorithm is based on Huygens' equivalence principle and nesting small algorithms within larger ones. Hence, the key element is to divide the computation into several stages and reduce the number of unknowns at each stage by a factor which is asymptotically equal to two. As a result, this algorithm is much faster than MOM. Since it has an asymptotic computational complexity of $N^{1.5}$, it has a great potential

for scattering computations of very large objects. Also, the structure of the algorithm lends itself for massively parallel processor computers. A pressing need for research is to reduce the proportionality constant in the CPU time analysis in (26). At this point, the proportionality constant is large.

Other algorithms with possibly less complexity are possible, for example, the CG-FFT algorithm [30], with a complexity of $O(N_{it}N \log N)$, where N_{it} is the number of iterations in the conjugate gradient method. However, our solution is valid for all incident waves while the CG-type algorithms provide a solution for only one incident wave. The method described here has the same complexity as those in the nested dissection ordering method [22]–[24] for solving differential equations. However, an integral equation solution has the advantage of satisfying the radiation condition.

As a final note, even though we have presented our theory on a rectilinear grid, the algorithm is not restricted to such grids.

X. APPENDIX

Reproduced here are the explicit expressions for $\bar{\mathbf{h}}_{ij}^{(o)}$ and $\bar{\mathbf{h}}_{ij}^{(i)}$ in two dimensions. We can show that

$$\bar{\mathbf{h}}_{ij}^{(o)} = \bar{\mathbf{c}}_i \cdot \bar{\boldsymbol{\alpha}}_{ij}, \quad (\text{A1})$$

where

$$\bar{\mathbf{c}}_i = \begin{bmatrix} 0 & (n_x + in_y) & 0 \\ (-n_x + in_y) & 0 & (n_x + in_y) \\ 0 & (-n_x + in_y) & 0 \end{bmatrix}. \quad (\text{A2})$$

More explicitly, it is

$$\begin{aligned} [\bar{\mathbf{h}}_{ij}^{(o)}]_{-1,m} &= \frac{ik\Delta}{8}(n_x + in_y)\psi_m(\mathbf{r}_i - \mathbf{r}'_j), \\ [\bar{\mathbf{h}}_{ij}^{(o)}]_{0,m} &= \frac{ik\Delta}{8}[(n_x + in_y)\psi_{m-1}(\mathbf{r}_i - \mathbf{r}'_j) \\ &\quad + (-n_x + in_y)\psi_{m+1}(\mathbf{r}_i - \mathbf{r}'_j)], \\ [\bar{\mathbf{h}}_{ij}^{(o)}]_{1,m} &= \frac{ik\Delta}{8}(-n_x + in_y)\psi_m(\mathbf{r}_i - \mathbf{r}'_j), \end{aligned} \quad (\text{A3})$$

where $m = -1, 0, 1$ and, hence, $\bar{\mathbf{h}}_{ij}^{(o)}$ is a 3×3 matrix. In the above, n_x and n_y are x and y components of the outward normal at $\mathbf{r}_i$. It can also be shown that

$$\bar{\mathbf{h}}_{ij}^{(i)} = \bar{\boldsymbol{\alpha}}_{ij} \cdot \bar{\mathbf{c}}_j, \quad (\text{A4})$$

or, more explicitly,

$$\begin{aligned} [\bar{\mathbf{h}}_{ij}^{(i)}]_{m,-1} &= \frac{ik\Delta}{8}(-n_x + in_y)\psi_{-m}(\mathbf{r}_i - \mathbf{r}'_j), \\ [\bar{\mathbf{h}}_{ij}^{(i)}]_{m,0} &= \frac{ik\Delta}{8}[(n_x + in_y)\psi_{-m-1}(\mathbf{r}_i - \mathbf{r}'_j) \\ &\quad + (-n_x + in_y)\psi_{-m+1}(\mathbf{r}_i - \mathbf{r}'_j)], \\ [\bar{\mathbf{h}}_{ij}^{(i)}]_{m,1} &= \frac{ik\Delta}{8}(n_x + in_y)\psi_{-m}(\mathbf{r}_i - \mathbf{r}'_j), \end{aligned} \quad (\text{A5})$$

where now, n_x and n_y correspond to the inward normal at $\mathbf{r}'_j$. In the above,

$$\psi_m(\mathbf{r}_i - \mathbf{r}'_j) = H_m^{(1)}(k|\mathbf{r}_i - \mathbf{r}'_j|)e^{im\phi_{ji}}, \quad (\text{A6})$$

where ϕ_{ji} is the angle the vector $\mathbf{r}_i - \mathbf{r}'_j$ makes with the x axis. Also [12]

$$\begin{aligned} [\bar{\boldsymbol{\alpha}}_{ij}]_{nm} &= H_{n-m}^{(1)}(k|\mathbf{r}_i - \mathbf{r}'_j|)e^{-i(n-m)\phi_{ji}} \\ &= H_{m-n}^{(1)}(k|\mathbf{r}_i - \mathbf{r}'_j|)e^{i(m-n)\phi_{ji}} \\ &= \psi_{m-n}(\mathbf{r}_i - \mathbf{r}'_j). \end{aligned} \quad (\text{A7})$$

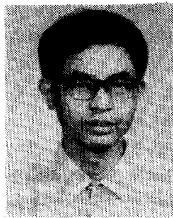
ACKNOWLEDGMENT

We would like to thank one of the reviewers for suggestions on improving the clarity of this manuscript.

REFERENCES

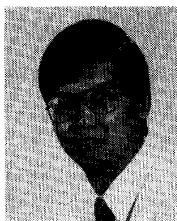
- [1] V. Roklin, "Rapid solution of integral equations of classical potential theory," *J. Comput. Phys.*, vol. 60, pp. 187–207, 1985.
- [2] L. Greengard and V. Roklin, "A fast algorithm for particle simulations," *J. Comput. Phys.*, vol. 73, pp. 325–348, 1987.
- [3] V. Roklin, "Rapid solution of integral equations of scattering theory in two dimensions," *J. Comput. Phys.*, vol. 86, pp. 414–439, 1990.
- [4] F. X. Canning, "Transformations that produce a sparse moment method matrix," *J. Electromagn. Waves Appl.*, vol. 4, pp. 903–913, 1990.
- [5] C. R. Anderson, "An implementation of the fast multipole method without multipoles," *SIAM J. Sci. Stat. Comput.*, vol. 13, no. 4, pp. 923–947, 1992.
- [6] W. C. Chew, "An N^2 algorithm for the multiple scattering solution of N scatterers," *Micro. Opt. Tech. Lett.*, vol. 2, no. 11, pp. 380–383, 1989.
- [7] W. C. Chew, J. Friedrich, and R. Geiger, "A multiple scattering solution for the effective permittivity of a sphere mixture," *IEEE Trans. Geoscience Remote Sensing*, vol. 28, no. 2, p. 207, 1990.
- [8] Y. M. Wang and W. C. Chew, "An efficient algorithm for solution of a scattering problem," *Micro. Opt. Tech. Lett.*, vol. 3, no. 3, pp. 102–106, March, 1990.
- [9] W. C. Chew and Y. M. Wang, "A fast algorithm for solution of a scattering problem using a recursive aggregate tau matrix method," *Micro. Opt. Tech. Lett.*, vol. 3, no. 5, pp. 164–169, May 1990.
- [10] Y. M. Wang and W. C. Chew, "Solution of large 2-D inhomogeneous scatterers for TM waves by fast recursive algorithm," *Micro. Opt. Tech. Lett.*, vol. 4, no. 3, Mar. 1991.
- [11] J. H. Lin and W. C. Chew, "Scattering solution of oblong, dielectric-coated, metallic scatterers by the recursive algorithm," *Electron. Lett.*, vol. 28, no. 2, pp. 185–186, January 1992.
- [12] W. C. Chew *et al.*, "A generalized recursive algorithm for wave-scattering solutions in two dimensions," *IEEE Trans. Microwave Theory Tech.*, vol. 40, pp. 716–723, Apr. 1992.
- [13] W. C. Chew, Y. M. Wang and L. Gurel, "A recursive algorithm for wavescattering using windowed addition theory," *J. Electromagn. Waves Appl.*, vol. 6, no. 11, pp. 1537–1560, 1992.
- [14] L. Gurel and W. C. Chew, "A recursive T-matrix algorithm for strips and patches," *Radio Sci.*, vol. 27, no. 3, pp. 387–401, 1992.
- [15] L. Gurel and W. C. Chew, "Scattering solution of three-dimensional, array of patches using the recursive T-matrix algorithms," *Micro. Guided Wave Lett.*, vol. 2, no. 5, 1992.
- [16] L. Gurel and W. C. Chew, "Recursive T-matrix algorithms for the solution of electromagnetic scattering from strip and patch geometries," *IEEE Trans. Antennas Propagat.*, vol. 41, pp. 91–99, Jan. 1993.
- [17] W. C. Chew and C. C. Lu, "NEPAL—An $N^{1.5}$ algorithm for solving the volume integral equation," *IEEE Antennas Propagat. Soc. Int. Symp. Dig.* (Chicago) July 18–25, 1992, pp. 184–187.
- [18] M. Born and E. Wolf, *Principles of Optics*. New York: Pergamon Press, 1970.
- [19] J. A. Kong, *Theory of Electromagnetic Waves*. New York: Wiley, 1975.
- [20] W. C. Chew, *Waves and Fields in Inhomogeneous Media*. New York: Van Nostrand Reinhold, 1990.
- [21] J. W. Cooley and J. W. Tukey, "An algorithm for the machine computation of complex Fourier series," *Math. Comp.*, vol. 19, pp. 297–301, 1965.
- [22] A. George, "Nested dissection of a regular finite element mesh," *SIAM J. Numer. Anal.*, vol. 10, pp. 345–363, 1973.
- [23] A. George and J. W. Liu, *Computer solution of large sparse positive definite system*. Englewood Cliffs, NJ: Prentice-Hall, 1981.
- [24] O. Axelsson and V. A. Barker, *Finite Element Solution of Boundary Value Problems: Theory and Computation*. New York: Academic Press, 1984.

- [25] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover Publications, 1965.
- [26] L. Gurel and W. C. Chew, "Recursive algorithms for the scattering by N disks or strips," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 507-515, Apr. 1990.
- [27] B. Peterson and S. Ström, "T-matrix for electromagnetic scattering from an arbitrary number of scatterers and representation of $E(3)$," *Phys. Rev.*, D8, pp. 3661-3678, 1973.
- [28] R. Harrington, *Field Computation by Moment Method*. Malabar, FL, 1983.
- [29] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes, the Art of Scientific Computing*. New York: Cambridge University Press, 1986.
- [30] D. T. Borup and O. P. Gandhi, "Fast-Fourier transform method for calculation of SAR distributions in finely discretized inhomogeneous models of biological bodies," *IEEE Trans. Microwave Theory Tech.*, vol. 32, pp. 355-360, Apr. 1984.



Cai-Cheng Lu was born in Hubei, China, on October 12, 1962. He received B.S. and M.S. degrees, both in electrical engineering, from the Beijing University of Aeronautics and Astronautics, China, in 1983 and 1986, respectively. Currently, he is pursuing the Ph.D. degree at the University of Illinois at Urbana-Champaign.

From 1986 to 1990, he was with the Department of Electronic Engineering, University of Aeronautics and Astronautics, China. From 1991 to 1992, he was a visiting scholar at the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign. His interests are in electromagnetic scattering, antennas, and wave propagation.



Weng Cho Chew (S'79-M'80-SM'86-F'93) was born on June 9, 1953 in Malaysia. He received the B.S. degree in 1976, both the M.S. and Engineer's degrees in 1978, and the Ph.D. degree in 1980, all in electrical engineering from the Massachusetts Institute of Technology, Cambridge.

His recent research interest has been in the area of wave propagation, scattering, inverse scattering, and fast algorithms related to inhomogeneous media for geophysical subsurface sensing and nondestructive testing applications. Previously, he has also analyzed electrochemical effects and dielectric properties of composite materials, microwave and optical waveguides, and microstrip antennas. From 1981 to 1985, he was with Schlumberger-Doll Research in Ridgefield, CT. There he was a program leader and later a department manager. From 1985 to 1990, he was an associate professor with the University of Illinois. He currently is a professor there and teaches graduate courses on waves and fields in inhomogeneous media and theory of microwave and optical waveguides and supervises a research program. His name is often listed in the **List of Excellent Instructors** on campus. He is author of the book *Waves and Fields in Inhomogeneous Media*, and he has published over 100 scientific journal articles and presented over 100 conference papers.

Dr. Chew is a member of Eta Kappa Nu, Tau Beta Pi, URSI Commissions B and F, and an active member with the Society of Exploration Geophysics. He was an NSF Presidential Young Investigator for 1986. He was also an AdCom member and is presently an associate editor of the *TRANSACTIONS* of the IEEE Geoscience and Remote Sensing Society. In addition he is an associate editor of the *International Journal of Imaging Systems and Technology* and the *Asia Pacific Engineering Journal* and has been a Guest Editor of *Radio Science* and the *International Journal of Imaging Systems and Technology*. He serves as associate director of the Advanced Construction Technology Center and the Electromagnetics Laboratory at the University of Illinois.