Published in final edited form as:

J Comput Phys. 2011 May 10; 230(10): 3656–3667. doi:10.1016/j.jcp.2011.02.016.

Reduced-Rank Approximations to the Far-Field Transform in the Gridded Fast Multipole Method

Andrew J. Hesford^{a,*} and Robert C. Waag^{a,b}

^a Department of Electrical and Computer Engineering, University of Rochester, Rochester NY 14642-8648 USA

^b Department of Imaging Sciences, University of Rochester, Rochester NY 14642-8648 USA

Abstract

The fast multipole method (FMM) has been shown to have a reduced computational dependence on the size of finest-level groups of elements when the elements are positioned on a regular grid and FFT convolution is used to represent neighboring interactions. However, transformations between plane-wave expansions used for FMM interactions and pressure distributions used for neighboring interactions remain significant contributors to the cost of FMM computations when finest-level groups are large. The transformation operators, which are forward and inverse Fourier transforms with the wave space confined to the unit sphere, are smooth and well approximated using reduced-rank decompositions that further reduce the computational dependence of the FMM on finest-level group size. The adaptive cross approximation (ACA) is selected to represent the forward and adjoint far-field transformation operators required by the FMM. However, the actual error of the ACA is found to be greater than that predicted using traditional estimates, and the ACA generally performs worse than the approximation resulting from a truncated singular-value decomposition (SVD). To overcome these issues while avoiding the cost of a full-scale SVD, the ACA is employed with more stringent accuracy demands and recompressed using a reduced, truncated SVD. The results show a greatly reduced approximation error that performs comparably to the full-scale truncated SVD without degrading the asymptotic computational efficiency associated with ACA matrix assembly.

Keywords

Fast solvers; iterative methods; acoustic scattering; fast multipole method; adaptive cross approximation; moment methods

1. Introduction

Fast scattering solutions have a broad range of applications. Estimation and correction of aberration requires simulation of propagation of ultrasound pulses through aberrating media [1–3]. Iterative inverse scattering algorithms such as the distorted Born iterative method [4–7] and the eigenfunction method [8–10] require repeated computation of the forward scattering problem for successive guesses of a reconstructed medium.

© 2011 Elsevier Inc. All rights reserved.

*Corresponding author: hesford@ece.rochester.edu (Andrew J. Hesford).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

The fast multipole method (FMM) [11–14] satisfies these needs by providing efficient solutions of integral equations that describe electromagnetic or acoustic scattering in two or three dimensions. The FMM recursively subdivides scattering media into groups of elements and computes the product of a test vector and the Green's function matrix in a time that is $O(N)$ when the medium is composed of N volume-filling elements. Redundant information in the FMM is discarded by representing interactions using band-limited wave expansions.

Representation of scattering media as sets of regular, gridded arrangements of elements provides a number of advantages. In inverse scattering applications, the scattering structure is not known *a priori*, rendering more sophisticated geometric meshes useless. Furthermore, scattering models developed with other techniques, such as magnetic resonance or b-scan imaging, are often inherently defined on a regular grid of voxels. Regular grids also allow the FMM to take advantage of FFT convolution to represent near-field interactions [15]. This technique, which exploits concepts originally developed for variants of the conjugate gradient-fast Fourier transform (CG-FFT) method [16–19], reduces the dependence of the FMM on the size of groups of scattering elements at the finest level. This is an advantageous hybrid because the FMM offers better asymptotic scaling than CG-FFT methods. Furthermore, unlike CG-FFT methods, the FMM can handle scattering by sparse arrangements of scatterers without the need to mesh the entire volume enclosing the scatterers.

Even with FFT convolution of near-field interactions, the FMM still requires transformations between the plane-wave expansions used in far-field interactions and the pressure distributions used for near-field interactions. When finest-level groups have edge lengths that approach a wavelength and contain hundreds or thousands of scattering elements, these transformations become the costliest portion of the FMM. Reducing the complexity of applying these transformations further relaxes the restriction on the size of finest-level FMM groups and provides additional opportunities to optimize the forward solution.

The adaptive cross approximation (ACA) [20–22] was developed to efficiently compute approximate, reduced-rank decompositions of matrices representing integral equations with smooth kernels. The method is purely algebraic and does not require knowledge of the full matrix to construct the approximation. The ACA has been applied to scattering problems by applying a recursive grouping of the scattering region in the same fashion as FMM and using ACA matrices to represent interactions between well-separated groups at the coarsest possible subdivision. Near-field interactions are represented, as in the fast multipole method, using dense matrices. However, because the ACA is purely algebraic and fails to directly exploit the integral scattering equation and its Green's function, approximate interaction matrices for mid-frequency scattering often have excessively high rank. The FMM, in contrast, employs an explicit decomposition of the Green's function to better control the bandwidth of far-field expansions, and hence the rank of the scattering operator, to reduce the amount of computation required for a certain accuracy requirement. In addition, the cost of building a hierarchical ACA scattering matrix scales with the square of the number of elements in the worst case [21]. Nevertheless, when used in concert with the FMM, the ACA can be an important part of efficient forward solution methods.

Previous studies have illustrated the efficiency of the FMM when combined with regular arrangements of scattering elements and FFT convolution of near-field interactions [15, 23]. That work is extended in this paper by employing the adaptive cross approximation to represent the far-field transformations that convert between far-field plane-wave expansions and local pressure distributions in the finest-level groups of the FMM. Because these operations are essentially Fourier transforms restricted to the unit sphere in wave space, the

transformations are smooth and well approximated by the ACA. The effective rank of the approximate far-field transform for an FMM group scales with the bandwidth of the plane-wave expansions associated with the group, which in turn scales approximately with the radius of a sphere enclosing the group. In contrast, the number of samples of the local pressure distribution scales with the volume of the group for a fixed sampling density. As a result, as the number of samples of local pressure distributions within finest-level groups are increased, low-rank approximations to far-field transforms become increasingly compressed. However, trials of the ACA show that it underestimates approximation error when using traditional convergence tests. To overcome this limitation, the far-field transformation matrices are approximated to lower error tolerances with ACA and then “recompressed” [24] using a reduced singular-value decomposition that provides error and rank characteristics comparable to the optimum reduced-rank approximation provided by truncated singular-value decomposition.

2. Theory

The scattering formulation employed herein is identical to that developed in Ref. [15], but is summarized below for convenient reference. Of interest is the solution of three-dimensional scattering of acoustic waves by a bounded, penetrable, inhomogeneous scattering region with arbitrarily variable sound speed $c(\mathbf{r})$, attenuation $\alpha(\mathbf{r})$, and density $\rho(\mathbf{r})$, where $\mathbf{r}$ is a three-dimensional coordinate vector. The region is embedded in an infinite, homogeneous background medium and is characterized by a complex wave number [10, 25, 26]

$$k(\mathbf{r}) = \frac{\omega}{c(\mathbf{r})} + i\alpha(\mathbf{r}), \quad (1)$$

where ω is the radian frequency of the time-harmonic pressure fields. A time dependence of $e^{-i\omega t}$ is assumed and suppressed to yield the acoustic wave equation [27]

$$\rho(\mathbf{r}) \nabla \cdot [\rho^{-1}(\mathbf{r}) \nabla p(\mathbf{r})] + k^2(\mathbf{r}) p(\mathbf{r}) = -S(\mathbf{r}), \quad (2)$$

in which $p(\mathbf{r})$ is the total acoustic pressure and S is an arbitrary acoustic source distribution.

To simplify Eq. (2), the change of variable $p(\mathbf{r}) = \rho_n^{1/2}(\mathbf{r}) f(\mathbf{r})$ is made, where $\rho_n = \rho/\rho_0$ is the ratio of the variable density to the homogeneous background density ρ_0 . Under this change of variable, Eq. (2) becomes the Helmholtz equation for the pseudo-pressure $f(\mathbf{r})$ [28, 29]:

$$\nabla^2 f(\mathbf{r}) + k_0^2 f(\mathbf{r}) = -\rho_n^{-1/2}(\mathbf{r}) S(\mathbf{r}) - O(\mathbf{r}) f(\mathbf{r}). \quad (3)$$

The object contrast or scattering potential is

$$O(\mathbf{r}) = k^2(\mathbf{r}) - k_0^2 - \rho_n^{1/2}(\mathbf{r}) \nabla^2 \rho_n^{-1/2}(\mathbf{r}). \quad (4)$$

The wave number k_0 corresponds to the homogeneous background with sound speed c_0 and attenuation α_0 .

When ρ is continuous, the Helmholtz equation (3) is readily inverted by using the homogeneous Green's function

$$g_0(\mathbf{r}, \mathbf{r}') = \frac{e^{ik_0|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \quad (5)$$

to yield an integral equation of scattering

$$f(\mathbf{r}) - \int_D d\mathbf{r}' g_0(\mathbf{r}, \mathbf{r}') O(\mathbf{r}') f(\mathbf{r}') = f^i(\mathbf{r}), \quad (6)$$

in which the incident field $f^i(\mathbf{r})$ satisfies the contrast-free Helmholtz equation

$$[\nabla^2 + k_0^2] f^i(\mathbf{r}) = -\rho_n^{-1/2}(\mathbf{r}) S(\mathbf{r}). \quad (7)$$

The integral expression in Eq. (6) is the scattered pseudo-pressure and is equal to the scattered pressure wherever $\rho_n = 1$.

It is important to note that, when ρ_n is discontinuous, solutions to Eq. (6) may correspond with acoustic pressures that are discontinuous or have a discontinuous derivative. This violates the physical boundary conditions governing wave propagation. These boundary conditions require that the pressure and its normal derivative remain continuous across media interfaces. However, approximations of the term $\rho_n^{-1/2} \nabla^2 \rho_n^{-1/2}$ appearing in Eq. (4) using finite differences on rectilinear grids with nonzero spacing inherently approximate discontinuous density profiles as C^1 -continuous functions. Therefore, the underlying issue is whether discontinuous density variations can be suitably approximated by continuous density variations. For problems in which the approximation is unsuitable, it may be necessary to modify the scattering formulation to correctly model the behavior as done in, *e.g.*, Ref. [30].

3. Methods

For numerical solution of Eq. (6), the scattering domain D is subdivided into a collection $\{c_j: 0 \leq j < N\}$ of N disjoint, cubic cells. Each cell c_j has a volume Δ_j and is associated with a constant contrast value O_j and a constant field value f_j such that

$$O(\mathbf{r}) = \sum_{j=0}^{N-1} O_j \chi_j(\mathbf{r}), \quad (8a)$$

$$f(\mathbf{r}) = \sum_{j=0}^{N-1} f_j \chi_j(\mathbf{r}), \quad (8b)$$

where χ_j is an arbitrary basis function with support c_j . As in Ref. [15], χ_j is here taken to be the characteristic function, or pulse basis function, of c_j . Together with Galerkin testing, the expansions in Eq. (8) allow representation of Eq. (6) as a discrete matrix equation

$$f - G \tilde{f} = f^i, \quad (9)$$

where the column vectors $f = [\Delta_j f_j]$, $\tilde{f} = [O_j f_j]$; the column vector $f^i = [f_j^i]$ contains entries

$$f_j^i = \int_D d\mathbf{r} \chi_j(\mathbf{r}) f^i(\mathbf{r}) \quad (10)$$

and the Green's matrix $G = [G_{ij}]$ contains entries

$$G_{ij} = \int_D d\mathbf{r} \chi_i(\mathbf{r}) \int_D d\mathbf{r}' g_0(\mathbf{r}, \mathbf{r}') \chi_j(\mathbf{r}'). \quad (11)$$

Iterative solution methods such as the generalized minimal residual method (GMRES) [31] or the stabilized biconjugate gradient method (BiCG-STAB) [32] may be employed in concert with the fast multipole method to solve Eq. (9) efficiently without the need to explicitly construct the Green's matrix (11) [13]. Such solutions are known to proceed in $O(N)$ time and with $O(N)$ storage.

3.1. The Fast Multipole Method

The FMM and its implementation are discussed in detail in Ref. [13]. At the heart of the method, inter-actions between source elements and target elements are represented as the composition of five operations:

1. Fields radiated by sources are *converted* by a far-field transform into outgoing wave expansions.
2. Neighboring sources are recursively grouped and their expansions are *aggregated* into outgoing expansions of successively higher bandwidth.
3. Outgoing expansions centered on a source group are *translated* into incoming wave expansions centered on a target group.
4. Incoming expansions are recursively *distributed* to target elements within groups.
5. Incoming expansions are converted by an adjoint far-field transform into fields observed at each target element.

In the Fourier domain, the outgoing and incoming wave expansions are plane-wave distributions. The aggregation, distribution and translation operators for such expansions are diagonal and efficiently applied. The increasing bandwidth of successively aggregated outgoing expansions is offset by a reduction in the number of field expansions that must be computed. Likewise, the increasing number of successively distributed incoming expansions is offset by the decreasing bandwidth of each expansion.

Translation of outgoing wave expansions to incoming expansions requires that the source and target group be sufficiently separated. Therefore, application of the FMM to Eq. (9) results in an expression of the form

$$f - [G_F + G_N] \tilde{f} = f^i, \quad (12)$$

where G_F represents far-field interactions that may be computed using the FMM and G_N represents near-field interactions that are computed directly using Eq. (11) among neighboring finest-level groups. In typical FMM implementations, the finest-level groups are kept small to reduce the impact of high-cost near-field interactions. However, as described in Ref. [15], FFT convolution may be employed when scattering elements are defined on a regular grid to reduce significantly the cost of near-field interactions. In such cases, it is the cost of evaluating far-field transforms (and their adjoints) of each finest-level group that dominates the cost of the FMM for large group sizes.

Far-field transforms represent outgoing plane-wave expansions for an FMM group as a function of the local pressure variation defined using basis functions within the group; adjoint far-field transforms represent an induced local pressure variation on testing functions within an FMM group as a function of incoming plane-wave expansions from radiation sources far from the group. For the described scattering problem, an arbitrary basis or testing function χ_j is associated with corresponding outgoing or incoming plane-wave expansion given, respectively, by [13]

$$F_j(\mathbf{c}_J, \hat{s}) = \int_D d\mathbf{r} e^{ik_0 \hat{s} \cdot (\mathbf{c}_J - \mathbf{r})} \chi_j(\mathbf{r}), \quad (13a)$$

$$R_j(\mathbf{c}_J, \hat{s}) = \int_D d\mathbf{r} e^{ik_0 \hat{s} \cdot (\mathbf{r} - \mathbf{c}_J)} \chi_j(\mathbf{r}). \quad (13b)$$

The outgoing expansion (13a) describes the far-field radiation pattern, in a direction $\hat{s}$ relative to an arbitrary center $\mathbf{c}_J$, of a unit pseudo-pressure applied to cell c_j . In a similar fashion, the incoming expansion (13b) performs a “focusing” onto the cell c_j of a plane wave incident from a direction $\hat{s}$ relative to an arbitrary center $\mathbf{c}_J$.

Consider a group $J = \{c_j; 0 \leq j < M\}$ of M scattering elements with center $\mathbf{c}_J$. Each cell c_j has an associated contrast coefficient O_j and a pseudo-pressure coefficient f_j . The outgoing plane-wave expansion β_J of the group is given by the far-field transform

$$\beta_J(\hat{s}) = k_0 \sum_{j=0}^M F_j(\mathbf{c}_J, \hat{s}) O_j f_j. \quad (14a)$$

The function β_J describes the complex amplitudes of plane waves traveling outward in directions $\hat{s}$ from $\mathbf{c}_J$. Conversely, a set of plane waves converging on the group J from directions $\hat{s}$ relative to $\mathbf{c}_J$ and with a complex amplitude profile γ_J induces a pressure on cell c_j described by the adjoint far-field transform

$$f_j = \int_{\Omega} d\hat{s} R_j(\mathbf{c}_J, \hat{s}) \gamma_J(\hat{s}), \quad (14b)$$

in which Ω is the unit sphere. When the plane-wave coefficients β_J and γ_J are sampled at discrete locations on the unit sphere, the integral in Eq. (14b) becomes a sum and the transforms (14) may be represented as matrix equations. The far-field transform is represented by a matrix, F , whose entries $F_{ij} = k_0 F_j(\mathbf{c}_J, \hat{s}_i)$. The adjoint far-field transform is represented by a matrix, R , whose entries $R_{ij} = R_i(\mathbf{c}_J, \hat{s}_j) = F_{ji}^*$. Storage and multiplication of these matrices must be made more efficient when large FMM groups are desired.

3.2. Far-Field Transforms and the Adaptive Cross Approximation

The excess bandwidth formula [13] prescribes an estimate of the harmonic bandwidth of the far-field pattern of fields radiated by groups of sources. If the group of sources is confined to a sphere with radius a in wavelengths, then the bandwidth L of the field radiated by the group is

$$L \approx ka + 1.8d_0^{2/3}(ka)^{1/3}, \quad (15)$$

in which $d_0 = -\log \varepsilon$ for a desired tolerance ε . Given an approximate bandwidth L , representations of plane-wave expansions of outgoing and incoming fields should be sampled at $O(L)$ polar locations and $O(L)$ azimuthal locations. In a cubic, gridded arrangement of N cubic scatterers, $a = O(N^{1/3})$, so the total number of samples of plane-wave expansions is $O(N^{2/3})$. Thus, the far-field transform (14a) may be represented by a matrix $F \in \mathbb{C}^{m \times n}$, where $m = O(N^{2/3})$ and $n = O(N)$. Similarly, the adjoint far-field transform (14b) may be represented as a matrix $R = F^\dagger \in \mathbb{C}^{n \times m}$, where $(\cdot)^\dagger$ represents the conjugate transpose of a matrix. Because the bandwidth of plane-wave expansions is approximately limited according to Eq. (15), reduced-rank approximate decompositions of the form

$$F \approx UV^\dagger, \quad (16)$$

with $U \in \mathbb{C}^{m \times k}$ and $V \in \mathbb{C}^{n \times k}$, have a rank $k = O(L) = O(N^{1/3})$. If such decompositions exist, the cost of applying a far-field transform or its adjoint will be reduced from $O(N^{5/3})$ to $O(N^{4/3})$.

The ACA [20–22] provides a method for computing an approximation of the form in Eq. (16). The ACA is an algebraic method that does not require full knowledge of the matrix being approximated and does not require replacement of the kernel of the underlying integral operator, as is done with the FMM. Conceptually, the ACA works by alternatively constructing columns to populate the column matrix U and the row matrix V . Columns for U and V are selected as the most significant of the remaining columns and rows, respectively, of the matrix F to be approximated; for this purpose, “most significant” means the column or row of F that contributes the largest element of the column of U or V that is currently being analyzed, neglecting elements contributed by previously analyzed columns or rows.

Algorithm 1 describes the ACA procedure for computing the columns u_k of the column matrix U and the columns v_k of the row matrix V . In the listing, $F_{i_k,:}$ represents the i_k -th row of F , while $F_{:,j_k}$ represents the j_k -th column of F . The ACA requires computation time that is $O(k^2[m + n])$ if the maximum rank of the decomposition is k . As implemented, the algorithm approximates the matrix F as the product $F = UV^T$. However, to maintain agreement with Eq. (16), the conjugate of the matrix V should be considered, such that $F = UV^\dagger$. The convergence criterion in Line 9 of Algorithm 1 is based on an estimate of the

approximation error [21]. The choice of 3 should reflect the fact that the actual approximation error, which is not revealed by the ACA algorithm, may be larger than that suggested by the convergence criterion.

3.3. Recompression of Adaptive Cross Matrices

The SVD of an arbitrary matrix $F \in \mathbb{C}^{m \times n}$ is given by

$$F = U \Sigma V^\dagger, \quad (17)$$

where $U \in \mathbb{C}^{m \times m}$ and $V \in \mathbb{C}^{n \times n}$ are unitary matrices and $\Sigma \in \mathbb{R}^{m \times n}$ is a diagonal matrix. The diagonal elements of Σ are called the singular values of the matrix F . By convention and to ensure uniqueness, the diagonal entries of Σ are generally assumed to be non-negative and arranged in order of decreasing magnitude. The truncated SVD is an approximation of rank k to the matrix F given by

$$F \approx F_k = \tilde{U} \tilde{\Sigma} \tilde{V}^\dagger, \quad (18)$$

in which $\tilde{U} \in \mathbb{C}^{m \times k}$ and $\tilde{V} \in \mathbb{C}^{n \times k}$ are unitary matrices consisting, respectively, of the first k columns of U and V , and $\tilde{\Sigma} \in \mathbb{R}^{k \times k}$ is a diagonal matrix whose entries are the largest k singular values of F . Of all matrices of rank k , the truncated SVD satisfies

$$F_k = \min_M \{ \|F - M\|_F : \text{rank } M = k \}, \quad (19)$$

where $\|\cdot\|$ is the Frobenius norm. Thus, adaptive cross approximations to far-field transform matrices will generally have greater error, in the sense of the Frobenius norm, than truncated singular-value decompositions of the transforms. Despite the lower error of truncated SVD approximations, the cubic scaling of the cost of computing an SVD is much worse than the $O(k^2[m+n])$ cost associated with computing a rank- k adaptive cross approximation to an $m \times n$ when $k \ll \min\{m, n\}$.

Recompression of ACA factorizations was first developed to eliminate redundant information by orthogonalizing the non-orthogonal row and column matrices resulting from the ACA [24]. If $F_k = UV^\dagger$ is an adaptive cross approximation of rank k to a matrix $F \in \mathbb{C}^{m \times n}$, the row and column matrices can each be orthogonalized using QR decompositions:

$$U = Q_u R_u, \quad (20a)$$

$$V = Q_v R_v. \quad (20b)$$

If the SVD of the $k \times k$ matrix product $R_u R_v^\dagger$ is given by

$$R_u R_v^\dagger = \widehat{U} \widehat{\Sigma} \widehat{V}^\dagger, \quad (21)$$

then the matrix F_k is given by the product

$$F_k = Q_u \widehat{U} \widehat{\Sigma} [Q_v \widehat{V}]^\dagger, \quad (22)$$

which is an SVD of F_k because the matrices $Q_u \widehat{U}$ and $Q_v \widehat{V}$ are orthogonal. This SVD is less costly to compute than a regular SVD of the matrix F_k because the $k \times k$ matrix $R_u R_v^\dagger$ has fewer elements than the $m \times n$ matrix F_k when $k < \min\{m, n\}$.

The redundancy resulting from the non-orthogonality of the row and column matrices computed using the ACA means that the error resulting in an approximation of fixed rank tends to have an error that is higher than estimated. Conversely, for a fixed approximation error, the ACA tends to overestimate the required rank of the decomposition. If an ACA factorization F_k of rank k is known to approximate a matrix F such that

$$\|F - F_k\|_F < \varepsilon_A \quad (23)$$

for some error tolerance ε_A , the reduced SVD in Eq. (21) may be truncated to $k' < k$ such that

$$\|R_u R_v^\dagger - \widehat{U}_{k'} \widehat{\Sigma}_{k'} \widehat{V}_{k'}^\dagger\| < \varepsilon_S, \quad (24)$$

where $\widehat{U}_{k'}$, $\widehat{\Sigma}_{k'}$ and $\widehat{V}_{k'}$ represent, respectively, the truncation of $\widehat{U}$, $\widehat{\Sigma}$ and $\widehat{V}$ to rank k' as described in Eq. (18). By the triangle inequality and the invariance of the Frobenius norm under transformation by unitary matrices, the truncated approximation satisfies

$$\|F - Q_u \widehat{U}_{k'} \widehat{\Sigma}_{k'} [Q_v \widehat{V}_{k'}]^\dagger\| < \varepsilon_A + \varepsilon_S. \quad (25)$$

Therefore, the rank of an adaptive cross approximation to a far-field transform can be reduced while ensuring a desired error tolerance ε by computing an initial approximation to a tolerance $\varepsilon_A < \varepsilon$ and then truncating a reduced SVD of the approximation to a tolerance ε_S suitably less than ε (e.g., $\varepsilon_S < \varepsilon - \varepsilon_A$). This technique avoids the higher cost of computing a full SVD of the far-field transform while providing a rank and approximation error comparable to that of the truncated SVD.

4. Numerical Results

Numerical studies were conducted on a scattering geometry designed to emulate a tissue-mimicking laboratory phantom. The geometry consisted of twelve scattering spheres embedded in a contrasting reference sphere of radius 24 mm with a sound speed of 1570 m/s, an absorption coefficient of 0.3 dB/(cm MHz), and a density of 970 kg/m³. This reference

sphere was immersed in an infinite, lossless background (water) with a sound speed of 1509 m/s and a density of 997 kg/m³. The interior spheres were each assigned properties to mimic one of the three human tissue types listed in Table 1. The tissue assignments and positions of the interior spheres, relative to the center of the reference sphere, are listed in Table 2. A diagram depicting the sphere arrangement is shown in Fig. 1.

For numerical modeling, the phantom is represented on a cubic domain whose center coincides with the center of the reference sphere and that has edge length 16λ , where λ is the wavelength in water. The domain is subdivided into elements with edge length 0.1λ , resulting in a $160 \times 160 \times 160$ computational grid. The FMM hierarchy is arranged so that each finest-level group contains the same number of scattering elements placed in the same relative cubic arrangement. A single far-field transform and its adjoint can thus be reused for all groups. When the number of elements in a single group does not divide the total number of elements, the domain is extended so that all groups are completely filled.

4.1. Error and Rank of the Approximate Far-Field Transform

Studies of the error and rank of approximations the far-field transform depend only on two aspects of the scattering problem: the number of scattering elements in a finest-level group (which may be unique in each dimension) and the acoustic size of each element. Because the previously described phantom is modeled by cubic elements with an edge length of 0.1λ , the behavior of far-field transforms for groups of cells of this size is of particular interest. However, forward solutions are commonly modeled using unknowns with edge lengths between 0.1λ and 0.2λ . Therefore, studies of far-field transforms involving cells with edge lengths of 0.2λ are also presented as a worst-case scenario. As in *the* phantom model, the FMM groups are assumed to be cubic so that the total number of elements completely determines the arrangement of elements within the group.

Based on the excess bandwidth formula, the approximate rank of the far-field transform is expected to scale as $O(N^{1/3})$, where N is the total number of elements in a group. The ranks of approximations for various error tolerances produced by the ACA, the recompressed adaptive cross approximation (RACA), and truncated SVD are shown in Fig. 2 as a function of the number of elements in a finest-level group. The $O(N^{1/3})$ scaling of the approximate rank was observed for all desired tolerances when the edge length of the scattering elements was fixed at 0.1λ . With element edge lengths of 0.2λ , the rank appeared to scale worse than $O(N^{1/3})$, but better than the $O(N^{2/3})$ scaling describing the number of samples of plane-wave expansions. In this case, the second term of Eq. (15) was a significant contributor to the bandwidth for the group sizes under investigation. The bandwidth, and hence the approximate rank, is expected to grow like $O(N^{1/3})$ for group sizes larger than those investigated.

The tolerances reported in Fig. 2 were only estimates of the true error (in the sense of the Frobenius norm) of approximate far-field transforms. For ACA, the reported tolerance was the convergence parameter ε . For the truncated SVD, the rank was selected to eliminate all singular values less than ε . In the RACA, the ACA was performed with a convergence parameter 0.01ε , and the reduced SVD was truncated with a tolerance ε . The actual error of reduced-rank approximations to the far-field transform, with the rank determined as in Fig. 2, is shown in Fig. 3 as a function of the total number of unknowns in the finest-level group. Whether the edge length of the scattering elements was 0.1λ or 0.2λ , the truncated SVD resulted in an actual error very close to the desired error. For elements with edge lengths of 0.1λ , the RACA yielded an error very close to that for a truncated SVD; the ACA error was between 1 and 1.5 orders of magnitude higher. When the element edge length was increased to 0.2λ , the error associated with the ACA increased further. For these larger elements, the RACA is expected to perform similarly to truncated SVD provided that the initial ACA

decomposition is performed with sufficiently low error. An increase in element size results in, *ceteris paribus*, an increased actual error for an adaptive-cross approximation. Therefore, the previously mentioned convergence parameter 0.01ε for the initial ACA might need to be further reduced to obtain a desired RACA tolerance of ε . The determination of optimum convergence parameters for larger element sizes will require further investigation.

4.2. Error in the Scattering Operator

Studies of error in the entire scattering operator contributed by approximate far-field transforms are problem dependent, but provide insight into the effect of the approximate transforms in a realistic setting. An ideal, but impractical, study would compare numerical solutions obtained with approximate scattering operators to known, analytical solutions. One important limitation of such an analysis is the lack of exact solutions for geometries more complicated than very basic structures. Another important limitation is that, even if the exact solution were known, the iterative solution process could compound the error inherent in the scattering operator in complicated ways. For the purposes of this study, the measure of error was derived from the normalized residual. For a system of equations $Ax = b$, the normalized residual associated

with an approximate solution $\tilde{x}$ is

$$r_A = \frac{b - A\tilde{x}}{\|b\|}. \quad (26)$$

Because the scattering operators under consideration are distinguished by the method used to compute far-field transforms and the number of elements in a finest-level FMM group, residuals will here be denoted $r_{T,M}$, where T is one of “full”, “ACA”, or “RACA” to denote the full-rank, ACA, or recompressed ACA representation of the far-field transforms, respectively. Truncated SVD representations of the far-field transforms were not investigated because truncated SVD and RACA performed similarly. The index M represents the number of elements per dimension per finest-level group, with the total number of elements in the group being $N = M^3$.

The first measure of the error contributed by reduced-rank approximations to full scattering operators is shown in Fig. 4 as the difference $\|r_{T,M} - r_{\text{full},M}\|$ as a function of the size $N = M^3$ of the finest-level groups. The ACA was used to approximate far-field transforms with a convergence parameter $\varepsilon = 10^{-4}$ or $\varepsilon = 10^{-6}$. The RACA, with a tolerance of 10^{-4} , was obtained from an initial ACA with a convergence parameter $\varepsilon = 10^{-6}$ and a reduced, truncated SVD such that all singular values less than 10^{-4} were suppressed. In general, the errors plotted in Fig. 4 follow the same trends as the errors plotted in Fig. 3. The approximate solution was the same in all cases, and was chosen so that $\|r_{\text{full},20}\| < 10^{-6}$. Thus, the approximate solution can be assumed to be in the neighborhood of the actual solution.

An additional measure of error is shown in Fig. 5 as the norm $\|r_{T,M} - r_{\text{full},30}\|$ as a function of the size $N = M^3$ of finest-level groups. The residual $r_{\text{full},30}$ was chosen as a reference because the solution accuracy should be limited by the accuracy of integration in the far-field transform. By minimizing the number and influence of FMM interactions, $r_{\text{full},30}$ should be closest to the true residual for the examined range of group sizes. As the figure shows, both the RACA with a tolerance of 10^{-4} and the ACA with a convergence parameter $\varepsilon = 10^{-6}$ very closely follow the behavior of scattering operators with full-rank solutions.

For the ACA with a convergence parameter $\varepsilon = 10^{-4}$, error in the approximate far-field transform limits the accuracy of the solution.

4.3. Computational Complexity

Perhaps the primary motivation for selecting reduced-rank approximations to far-field transforms in the FMM is the availability of acceptably accurate solutions in less time, rather than the availability of highly accurate solutions. Thus, the time required to compute a matrix-vector product (MVP) using the FMM with approximate far-field transforms is of great interest. The CPU time required to compute an MVP using both ACA and full-rank far-field matrices is shown in Fig. 6. Also shown are the CPU times required for evaluations of forward and adjoint far-field transforms only and the CPU time required for evaluations of near-field interactions only. The time required to compute near-field interactions is the same whether full-rank or approximate far-field transforms are used.

As noted above, the scattering domain is extended when the $160 \times 160 \times 160$ grid is not uniformly divided by the size of finest-level FMM groups. Therefore, the number of unknowns in the problem is not constant, and the times required to compute matrix-vector products are influenced by this variation. The times reported in Fig. 6 have been normalized through division by the actual number of scattering elements and multiplication by 4, 096, 000. This eliminates from consideration the overall linear scaling of the FMM.

The cost of evaluating a single, full-rank far-field transform of N elements is $O(N^5/3)$, while the cost of evaluating a reduced-rank transform is $O(N^4/3)$. However, when the overall size of the scattering problem is fixed, the total number of groups varies like $O(1/N)$. Therefore, the total cost for evaluating full-rank far-field transforms is expected to scale like $O(N^{2/3})$, while the total cost for evaluating reduced-rank transforms scales like $O(N^{1/3})$. This behavior is confirmed by the plots in Fig. 6. Furthermore, the reduced computational dependence of the FMM of finest-level group size shifted the optimum group size from 5 elements per group per dimension to 10 elements per group per dimension and reduced the MVP time by approximately 33%.

5. Conclusion

An implementation of the FMM was presented that employs reduced-rank approximations to far-field transforms to reduce the dependence of the FMM on the size of finest-level groups of scatterers. While the ACA was found to underestimate error for a given convergence parameter and generally perform worse than truncated SVD, the recompressed ACA maintains the performance advantage of the ACA while providing error behavior comparable to the truncated SVD. The resulting reduced computational dependence facilitates the selection of an optimum group size to improve overall computation time by eliminating excessive FMM calculations.

Acknowledgments

Professor Weng C. Chew, of the University of Illinois and the University of Hong Kong, is thanked for the use of his ScaleME parallel FMM library [13, 33]. Jason C. Tillett and Jeffrey P. Astheimer, both of the University of Rochester, are thanked for helpful discussions, suggestions, and comments about material in this paper. This research was funded in part by NIH Grants EB 009692 and EB 010069 and the University of Rochester Diagnostic Ultrasound Research Laboratory Industrial Associates.

References

1. Tabei M, Mast TD, Waag RC. Simulation of ultrasonic focus aberration and correction through human tissue. *J Acoust Soc Am*. 2003; 113:1166–1176. [PubMed: 12597210]

2. Lacefield JC, Pilkington WC, Waag RC. Distributed aberrators for emulation of pulse distortion by abdominal wall. *Acoust Res Lett Online*. 2002; 3:47–52.
3. Mast TD, Hinkelman LM, Metlay LA, Orr MJ, Waag RC. Simulation of ultrasonic pulse propagation, distortion, and attenuation in the human chest wall. *J Acoust Soc Am*. 1999; 106:3665–3677. [PubMed: 10615705]
4. Devaney AJ, Oristaglio ML. Inversion procedure for inverse scattering within the distorted-wave Born approximation. *Phys Rev Lett*. 1983; 51:237–240.
5. Chew WC, Wang YM. Reconstruction of two-dimensional permittivity distribution using the distorted Born iterative method. *IEEE Trans Med Imag*. 1990; 9:218–225.
6. Lavarello R, Oelze M. Density imaging using inverse scattering. *J Acoust Soc Am*. 2009; 125:793–802. [PubMed: 19206857]
7. Hesford AJ, Chew WC. Fast inverse scattering solutions using the distorted Born iterative method and the multilevel fast multipole algorithm. *J Acoust Soc Am*. 2010; 128(2):679–690. [PubMed: 20707438]
8. Mast TD, Nachman AI, Waag RC. Focussing and imaging using eigenfunctions of the scattering operator. *J Acoust Soc Am*. 1997; 102:715–726. [PubMed: 9265750]
9. Lin F, Nachman AI, Waag RC. Quantitative imaging using a time-domain eigenfunction method. *J Acoust Soc Am*. 2000; 108:899–912. [PubMed: 11008794]
10. Waag RC, Lin F, Varslot TK, Astheimer JP. An eigenfunction method for reconstruction of large-scale and high-contrast objects. *IEEE Trans Ultrason, Ferroelectr Freq Control*. 2007; 54:1316–1332. [PubMed: 17718321]
11. Greengard L, Rokhlin V. A fast algorithm for particle simulations. *J Comput Phys*. 1987; 73:325–348.
12. Rokhlin V. Rapid solution of integral equations of scattering theory in two dimensions. *J Comput Phys*. 1990; 86:414–439.
13. Chew, WC.; Jin, J.; Michielssen, E.; Song, J., editors. *Fast and Efficient Algorithms in Computational Electromagnetics*. Artech House; Boston: 2001.
14. Michielssen E, Jin JM. Guest editorial for the special issue on large and multiscale computational electromagnetics. *IEEE Trans Antennas Propag*. 2008; 56(8):2146–2149.
15. Hesford AJ, Waag RC. The fast multipole method and Fourier convolution for the solution of acoustic scattering on regular volumetric grids. *J Comput Phys*. 2010; 229:8199–8210. [PubMed: 20835366]
16. Zhai H, Chen Q, Yuan Q, Sawaya K, Liang C. Analysis of large-scale periodic array antennas by CG-FFT combined with equivalent sub-array preconditioner. *IEICE Trans Commun*. 2006; E89-B: 922–928.
17. Cui TJ, Chew WC, Aydin AA, Zhang YH. Fast-forward solvers for the low-frequency detection of buried dielectric objects. *IEEE Trans Geosci Remote Sens*. 2003; 41:2026–2036.
18. Xu XM, Liu QH. The BCGS-FFT method for electromagnetic scattering from inhomogeneous objects in a planarly layered medium. *IEEE Antennas Wireless Propag Lett*. 2002; 1:77–80.
19. Cui TJ, Chew WC. Fast algorithm for electromagnetic scattering by buried 3-D dielectric objects of large size. *IEEE Trans Geosci Remote Sens*. 1999; 37:2597–2608.
20. Bebendorf M. Approximation of boundary element matrices. *Numer Math*. 2000; 86:565–589.
21. Zhao K, Vouvakis MN, Lee JF. The adaptive cross approximation algorithm for accelerated method of moments computations of EMC problems. *IEEE Trans Electromagn Compat*. 2005; 47:763–773.
22. Shaeffer J. Direct solve of electrically large integral equations for problem sizes to 1 M unknowns. *IEEE Trans Antennas Propag*. 2008; 56(8):2306–2313.
23. De Zaeytjij J, Bogaert I, Franchois A. An efficient hybrid MLFMA-FFT solver for the volume integral equation in case of sparse 3D inhomogeneous dielectric scatterers. *J Comput Phys*. 2008; 227:7052–7068.
24. Bebendorf M, Kunis S. Recompression techniques for adaptive cross approximation. *Journal of Integral Equations and Applications*. 2009; 21(3):331–357.

25. Nachman AI, Smith JF III, Waag RC. An equation for acoustic propagation in inhomogeneous media with relaxation losses. *J Acoust Soc Am*. 1990; 88(3):1584–1594.
26. Duck, FA. *Physical Properties of Tissue: A Comprehensive Reference Book*. Academic Press; London: 1990.
27. Morse, PM.; Ingard, KU. *Theoretical Acoustics*. McGraw-Hill; New York: 1968.
28. Johnson SA, Stenger F, Wilcox C, Ball J, Berggren MJ. Wave equations and inverse solutions for soft tissue. *Acoustic Imaging*. 1982; 11:409–424.
29. Pourjavid S, Tretiak OJ. Numerical solution of the direct scattering problem through the transformed wave equation. *J Acoust Soc Am*. 1992; 91:639–645. [PubMed: 1556316]
30. Martin PA. Acoustic scattering by inhomogeneous obstacles. *SIAM J Appl Math*. 2003; 64(1): 297–308.
31. Saad Y, Schultz MH. GMRES—a generalized minimal residual algorithm for solving nonsymmetric linear-systems. *SIAM J Sci and Stat Comput*. 1986; 7:856–869.
32. van der Vorst HA. Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems. *SIAM J Sci and Stat Comput*. 1992; 13(2):631–644.
33. Velamparambil S, Chew WC. Analysis and performance of a distributed memory multilevel fast multipole algorithm. *IEEE Trans Antennas Propag*. 2005; 53:2719–2727.

Figure 1.

An arrangement of twelve tissue-mimicking spheres used for numerical experiments. The enclosing reference sphere is not shown. The spheres are colored according to their composition as listed in Table 2. Red, green and blue represent skin, muscle, and fat, respectively.

Figure 2.

Comparison of the predicted rank for approximations to the far-field transform using the ACA, the RACA, and truncated SVD for groups containing a variable number of cells of edge length 0.1λ or 0.2λ . The desired tolerance of the approximations is shown along the right axis. For the 0.1λ edge length, the RACA and truncated SVD ranks coincided exactly. RACA was not investigated for edge lengths of 0.2λ .

Figure 3.

Actual error, in the sense of the Frobenius norm, between the actual far-field transform and the rank-limited approximations produced by the ACA, the RACA, and truncated SVD for cubic groups with a variable number of cells of edge length 0.1λ or 0.2λ .

Figure 4. Residual error of the scattering operator using ACA or RACA for far-field transforms for groups containing a variable number of elements, relative to the residual computed using full-rank far-field transforms.

Figure 5.

Residual error of the scattering operator using ACA, RACA and full-rank far-field transforms for groups with a variable number of elements with edge length 0.1λ , relative to the residual of the scattering operating using full far-field transforms with finest-level groups containing $30 \times 30 \times 30$ elements.

Figure 6.

Computing time for a single matrix-vector product (MVP) using FMM with a reduced-rank, approximate far-field transform using ACA with a tolerance of 10^{-4} , compared to the computing time for a single MVP when a full, dense far-field transform is employed. Also shown is the time for evaluating only forward and adjoint far-field transforms using full matrices or the ACA. The time for computing near-field interactions (NN) is identical in both implementations.

Table 1

Material properties of spheres designed to mimic human tissue.

Tissue	Sound speed m/s	Absorption dB/(cm MHz)	Density kg/m ³
Water	1509.0	0.00	997.0
Reference	1570.0	0.30	970.0
Fat	1478.0	0.52	950.0
Muscle	1547.0	0.91	1050.0
Skin	1613.0	1.61	1120.0

Table 2

Characteristics of the spheres in the tissue-mimicking phantom.

Radius (mm)	Center (mm)			Tissue
	x	y	z	
4.0	0.0	0.0	0.0	Fat
5.0	14.0	2.0	4.0	Skin
5.0	5.0	-10.0	-4.0	Fat
3.0	17.0	-7.0	0.0	Fat
7.8	-10.0	10.0	7.2	Muscle
7.8	5.0	12.0	-7.2	Muscle
5.0	14.0	12.0	3.0	Muscle
5.0	-5.0	-18.0	-3.0	Skin
2.5	7.5	-18.0	-2.0	Skin
1.5	-4.0	20.0	0.0	Skin
2.5	-18.0	4.0	2.0	Skin
9.1	-12.5	-5.0	-5.2	Muscle

Algorithm 1**The Adaptive Cross Approximation**

```

1:   $i_1 = 1$ 
2:   $\|F^{(0)}\| = 0$ 
3:  for  $k = 1$  to  $\min\{m, n\}$  do
4:     $v_k = F_{i_k,:} - \sum_{l=1}^{k-1} (u_l)_{i_k} v_l$ 
5:     $j_k = \operatorname{argmax}_{j \neq i_1, \dots, i_{k-1}} |(v_k)_j|$ 
6:     $v_k = v_k / (v_k)_{j_k}$ 
7:     $u_k = F_{:,j_k} - \sum_{l=1}^{k-1} (v_l)_{j_k} u_l$ 
8:     $\|F^{(k)}\|^2 = \|F^{(k-1)}\|^2 + \|u_k\|^2 \|v_k\|^2 + 2 \sum_{l=1}^{k-1} |u_l^T \cdot u_k| \cdot |v_l^T \cdot v_k|$ 
9:    if  $\|u_k\| \cdot \|v_k\| \leq \varepsilon$  or  $\|F^{(k)}\|$  then
10:      break
11:    end if
12:     $i_{k+1} = \operatorname{argmax}_{i \neq i_1, \dots, i_k} |(u_k)_i|$ 
13:  end for

```