Audioptimization: Goal-Based Acoustic Design

Michael Monks, Byong Mok Oh, and Julie Dorsey
Massachusetts Institute of Technology

We present a new inverse, interactive approach to acoustic design that applies optimization techniques to an acoustic simulation system.

Acoustic design is difficult because the human perception of sound depends on such things as decibel level, direction of propagation, and attenuation over time—none of which are tangible or visible. Traditionally, designers have built physical scale models and tested them visually and acoustically. For example, by coating the interiors of the models with reflective material and then shining lasers from various source positions, they assess the sight and sound lines of the audience in a hall. They also might attempt to measure acoustical qualities of a proposed environment by conducting acoustic tests on the model using sources and receivers scaled in both frequency and size. Even water models are used sometimes to visualize the acoustic wave propagation in a design. These traditional methods have proven inflexible, costly, and time-consuming to implement, and they have created some major acoustic failures.[1]

The advent of computer simulation and visualization techniques for acoustic design and analysis has yielded a variety of approaches for modeling acoustic performance.[2-4] Research in this area to date has mainly addressed the accuracy and speed of the simulation algorithms and offered effective visualizations or auralizations of the resulting multidimensional data. However, while these techniques certainly offer new insights into acoustic design, they fail to enhance the design process itself, which still involves a burdensome iterative process of trial and error.

Many complex, often-conflicting goals and constraints generally mark the design process. For instance, financial concerns might dictate a larger hall with increased seating capacity. This can have negative effects on the hall's acoustics, such as excessive reverberation and noticeable gaps between direct and reverberant sound. Fan-shaped halls bring the audience closer to the stage than other configurations, but they may fail to make the listener feel surrounded by the sound. The application of highly absorbent materials may reduce disturbing echoes, but they may also deaden the hall.

In many renovations, budgetary, aesthetic, or physical impediments limit modifications, compounding the difficulties confronting the designer. In addition, a hall might need to accommodate a wide range of performances, from lectures to symphonic music, each with different acoustic requirements. In short, a concert hall's acoustics depends on the designer's ability to balance many factors.

Here, we present an inverse, interactive acoustic design approach that helps a designer produce an architectural configuration that achieves a desired acoustic performance. For a new building, the system may suggest optimal configurations that would not otherwise be considered; for a hall with modifiable components or for a renovation project, it may assist in optimizing an existing configuration. Our system allows the designer to constrain changes to the environment and specify acoustic performance goals as a function of time. The constraints include the specification of a range of allowable materials as well as geometric modifications for surfaces in the hall. The designer also specifies goals for acoustic performance in space and time via high-level acoustic qualities such as decay time and sound level.

Using this information, the system performs a constrained optimization of surface material and geometric parameters for a subset of elements in the environment. The system operates at varying accuracy levels, offering trade-offs between time and quality. Visualization tools facilitate an intuitive assessment of the complex time-dependent nature of sound, and they provide a means to express desired performance. By using optimization routines within an interactive application, our system reveals complex acoustic properties and steers the design process toward the designer's goals.

Background

Our approach synthesizes and extends previous research in the areas of interactive design and optimization, acoustic simulation and visualization, and characterization of sound.

0272-1716/00/$10.00 © 2000 IEEE

Interactive design, optimization techniques

Today's CAD systems for lighting, acoustic, and other types of simulation-intensive design are based almost exclusively on direct methods: those that compute a solution from a complete description of an environment and relevant parameters. Such systems can be extremely useful in evaluating the performance of a given 3D environment. However, they involve a tedious specify-simulate-evaluate loop in which users are responsible for specifying input parameters and for evaluating the results. The computer is responsible only for computing and displaying the results of these simulations. The drawback of this method is especially noticeable if a costly simulation (for example, lighting or acoustic) is part of the loop. This makes interactive searching of the design space impossible.

Recently, Marks et al. developed a methodology, called design galleries, for searching large parameter spaces typical of design problems.[5] The system automatically generates a palette of widely spaced, distinct choices from which users may select the most appropriate one. While this technique allows users to explore a variety of different configurations, it is less useful in problems—such as acoustic design—for which the goals are known in advance or that have a large number of degrees of freedom.

An alternative approach to design considers the inverse problem—that is, allowing users to create a target and have the algorithm work backward to establish various parameters. In this division of labor, users must now specify objectives to achieve in a scene. The computer searches the design space, that is, it selects parameters optimally with respect to user-supplied objectives. Several lighting design and rendering systems have employed inverse design. For example, users can specify the location of highlights and shadows,[6] pixel intensities or surface radiance values,[7] or subjective impressions of illumination.[8] The computer then attempts to determine lighting or reflectance parameters that best match the given objectives using optimization techniques. Because sound is considerably more complex than light, an inverse approach appears to have even more potential in assisting acoustic designers. We build on previous inverse design systems by optimizing not only over materials but also over geometric parameters.

The role of optimization in a design system is to find the configuration in the feasible design space that best matches desired performance goals. The choice of an optimization technique depends on the nature of the design space and the types of constraints. Here, we formulate the acoustic design problem as a constrained, nonlinear optimization problem. The basic constrained optimization problem is to minimize the scalar quantity of an objective function of n system parameters while satisfying a set of constraints. Standard nonlinear optimization techniques use the gradient and curvature of the objective function to descend to a minimum or locally optimal configuration.[9] Our evaluation functions have multiple such minima and therefore require a global strategy. Hence, we first employ a global optimization technique—simulated annealing—to locate

The role of optimization in a design system is to find the configuration in the feasible design space that best matches desired performance goals. The choice of an optimization technique depends on the nature of the design space and the types of constraints.

a more globally optimal neighborhood. We then use the steepest descent algorithm to descend to the minimum.

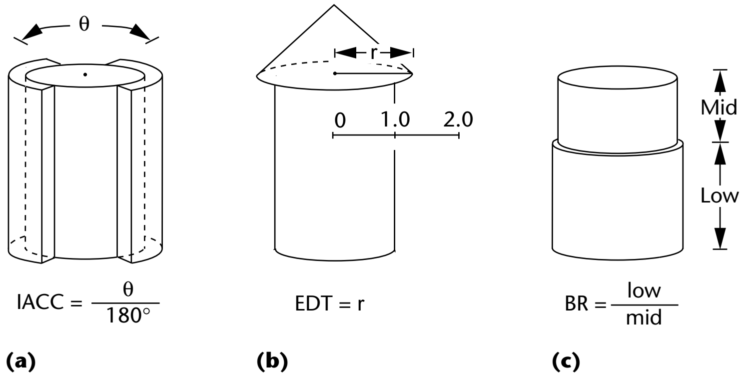
Acoustic simulation and visualization

Previous work in acoustic simulation[1] can be divided into five general categories: image source methods,[2] radiant exchange methods,[3,10] statistical methods, ray tracing,[11] and beam tracing.[3,12,13] A variety of hybrid simulation techniques typically approximate the sound field by modeling the early and late sound fields separately and combining the results.[3]

We employ such a hybrid simulation engine, which models the early sound with beam tracing and the late sound with a statistical approximation.[14] With this simulation approach, a receiver-independent preprocessing step determines the location of all virtual sources up to a specified reflection depth and calculates the subvolume of the room that each virtual source influences. The late sound is modeled by a statistical tail, a single source distributing the remaining energy to all locations equally as it decays. Once a receiver location is specified, we determine a list of valid virtual sources, each contributing sound energy to that receiver in the form of a discrete arrival. We can then calculate acoustic measures using the list of discrete arrivals and the precalculated statistical tail.

The simulation engine has several characteristics relevant to our work. First, since the geometric preprocessing step is independent of location and time, we can specify a receiver location at any stage of design and obtain its set of acoustic data merely by sampling the precalculated sound field. Second, the cost of modifying surface materials in the hall is a small fraction of that of modifying geometry. Changes in geometry invalidate a portion of the beam tree data structure and therefore require its reconstruction, while changes in materials require only that the energy contributions of affected beams be updated. Finally, it's possible to trade quality for speed by reducing the reflection depth of the beam tracing, culling beams with minimal contribution, and/or reducing sampling density. An important attribute of our system is that it can be used in conjunction with virtually any acoustic simulation algorithm.

Other work has been done in representing sound field



1 **Glyphs representing (a) IACC, (b) early decay time (EDT), and (c) bass ratio (BR).**

data with visualizations and auralizations. Stettner and Greenberg presented a set of 3D glyphs to convey graphically the behavior of sound within an enclosure.[15] The Bose Auditioner system provides auralizations from simulation data at listener positions within a modeled hall.[4] These auralizations approximate what the hall might sound like. Our system provides visualizations for a collection of acoustic measures that describe the character of the sound field as it varies in space and time within an environment. We use these visualizations both to analyze the behavior of a given design and to specify desirable performance goals interactively.

Characterization of sound

Traditionally, reverberation time and other early decay measurements were considered the primary evaluation parameters in acoustic design. However, in recent years, researchers have recognized the inadequacies of using these criteria alone and have introduced a variety of additional measures aimed at characterizing the subjective impression of human listeners.[1] In 1996 Beranek introduced an evaluation function that gives an overall acoustic rating by linearly combining six statistically independent objective acoustic measures.[16] This function builds on the work of Ando.[17] We employ Beranek's evaluation approach, known as the Objective Rating Method (ORM). We define the six acoustic measures and introduce visualization techniques used to evaluate them.

Interaural cross-correlation coefficient. IACC is a binaural measure of the correlation between the signals at the two ears of a listener. It characterizes how surrounded a listener feels by the sound within a hall. If the sound comes from directly in front of or behind the listener, it will arrive at both ears at the same time with complete correlation, producing no stereo effect. If it comes from another direction, the two signals will be out of phase and less correlated, giving the listener the desirable sensation of being enveloped by the sound.[17] The correlation values depend on the arrival direction of the wave with respect to the listener's orientation. Since the amplitude of sound decreases rapidly as it propagates, the sound waves that arrive the earliest generally have far greater effect on IACC.

We use a cylinder to represent a receiver in a performance space. The graphical icon we use to represent IACC, shown in Figure 1, is a shell located on the sides of each listener icon. The greater the degree of encirclement of the icon by the shell, the more desirable is the IACC value.

Early decay time. EDT measures the reverberation or liveliness of a hall. Musicians characterize a hall as "dead" or "live," depending on whether EDT is too low or high. The formal definition of EDT is the time it takes for the level of sound to drop 10 decibels from its initial level, which is then normalized for comparison to traditional measures of reverberation by multiplying the value by six. As Beranek suggests, we determine EDT by averaging the values of EDT for 500-Hz and 1,000-Hz sound pulses. The best values of EDT range between 2.0 and 2.3 seconds for concert halls. The icon we use to portray EDT at a receiver position is a cone in which the height is fixed and the radius is scaled according to the decay time (see Figure 1). For a value of 2.0 seconds, the cone is twice the width of the listener icon.

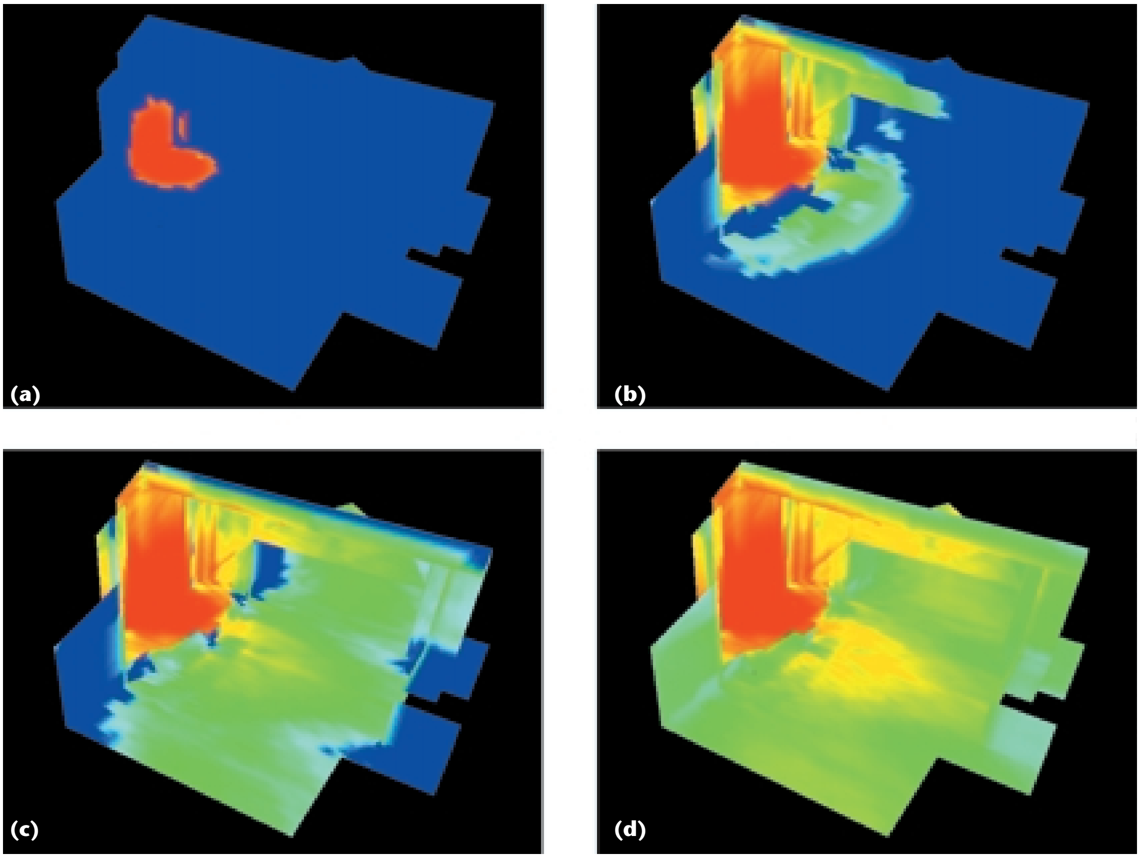
Bass ratio. BR measures how much sound comes from bass, reflecting the persistence of low-frequency energy relative to mid-frequency energy. Musicians refer to BR as the "warmth" of the sound. BR is defined as

$$BR = \frac{RT_{125} + RT_{250}}{RT_{500} + RT_{1000}}$$

RT_f is the frequency dependent reverberation time at an octave band centered at frequency f. RT is the time it takes for the sound level to drop from 5 dB to 35 dB below the initial level, which is then normalized for comparison to traditional measures of reverberation by multiplying by two. For example, for a 100-dB initial sound level, RT would be the time it takes to drop from 95 dB to 65 dB multiplied by the normalizing factor. The ideal value of BR ranges between 1.1 and 1.4 for concert halls.

The graphical icon we use to represent BR at a receiver position consists of a pair of stacked, concentric cylinders of slightly different widths (see Figure 1). The top cylinder represents the mid-frequency energy, from the 500-Hz and 1,000-Hz bands, and the bottom cylinder represents the lower frequency energy, from the 125-Hz and 250-Hz bands. The height of each cylinder represents the relative values in the ratio, assuming a constant combined height. A listener icon representing a desirable BR value of 1.25 would have the top of the lower cylinder just above the halfway mark, as shown in Figure 1. Like the other measures discussed so far, BR varies spatially throughout a hall.

Strength factor (G). This factor measures sound level, approximating a general perception of loudness of the sound in a space. For a given location within the hall, G is the ratio of the sound energy arriving at that location from a nondirectional source to the direct



2 Color indicates sound strength data on room surfaces, calculated for different integration intervals: (a) 10 ms, (b) 40 ms, (c) 80 ms, and (d) 120 ms.

sound energy measured at a distance of 10 meters from the same source. By using this ratio, the influence of source power is removed from the loudness calculation, allowing easy comparison of measured data across different halls. We average the values of G at 500 Hz and 1,000 Hz. The preferred values for G range between 4.0 dB and 5.5 dB for concert halls. In general, G is higher at locations closer to the source.

It is instructive to see how the sound level changes through time, as well as location. We perceive a reflected wave front as an echo—perceptibly separable from the initial sound—if it arrives more than 50 ms after the direct sound and if it is substantially stronger than its neighbors. The time distribution of sound also affects our perception of clarity. Two locations in a hall may have the same value of G, but if the energy arrives later with respect to the direct sound for one location than the other, speech will be less intelligible and music less crisp. We use color to indicate relative scalar values of the sound-level data, which is sampled at a fine mesh of points on surfaces as a function of time. By adjusting a slider, we can examine the accumulated sound-level data as a function of space and time. Figure 2 illustrates a representative time sequence.

Initial-time-delay gap (TI). This gap measures how large the hall sounds, quantifying the perception of intimacy the listener feels in a space. It depends purely on the hall's geometry, measuring the time delay between the arrival of the direct sound and that of the first reflected wave to reach the listener. To make comparisons among different halls, we record only a single value per hall, measured at a representative location in the center of the main seating area. It is best if TI does not exceed 20 ms.

Surface diffusivity index. SDI is a measure of the amount of sound diffusion caused by gross surface detail, or macroscopic roughness of surfaces within a hall. SDI is usually determined by inspection, and it correlates to the tonal quality of the sound in a hall. We compute the SDI index for the entire hall by summing the SDI assigned to each surface material, weighted by its area with respect to the total surface area of the hall. SDI can range between 0.0 and 1.0, with larger values representing more diffusion. The preferred value of SDI is 1.0. For example, plaster has a lower index than brick, which has a lower index than corrugated metal.

These six statistically orthogonal acoustic measures[16] form the basis for our analysis and optimization work. While two of the measures, SDI and TI, are single values representing the entire hall, we compute the others by averaging the values sampled at multiple spatial positions and, in the case of G, multiple points in time.

Inverse problem formulation

We phrase this problem more formally as follows: given a description of a set of desired measures for acoustic performance, determine the material properties and geometric configuration that will most closely match the target. To formulate the acoustic design process as a constrained optimization problem, we require a specification of (1) the optimization variables that express how a hall may be modified, (2) the constraints that must be satisfied, and (3) the objective function.

Optimization variables

In a typical acoustic simulation system, the goal is to compute the sound field in a scene assuming a sound source and a description of the geometry and materials. The measures just described are the unknowns, which are computed in terms of static material properties and geometry. In the optimization problem, material and geometric properties are no longer fixed but are treated as variables.

A hall consists of a collection of polygons, subsets of which may be grouped into geometric components. Components are a convenient and natural way to represent entities such as balconies, reflectors, and so on. Each component can have associated with it a set of allowable linear geometric transformations and a set of acceptable materials. Each translation, rotation, or scaling of a component represents a geometry variable; a set of possible materials associated with a component is a material variable.

Constraints

There are two types of constraints. Geometry constraints are user-specified upper and lower bounds placed on each component's transformation. Each transformation variable represents a single degree of freedom: translation along a vector, rotation about a vector, or scaling about a point or along a vector. The allowable range of each transformation constraint requires the component to remain within the specified bounds. For example, $T_{low} \le T_i \le T_{high}$ requires the transformation i to remain within the bounds of T_{low} and T_{high}.

Material constraints are user-specified sets of allowable materials assigned to a given component. This subset of materials is selected from a library of materials provided by the system. For example, {*plaster,concrete,fiberglass*} is a set of materials for a component. The library is built such that all material properties are physically sensible.

Objective function

Acoustic design problems are typically underconstrained. Hence, an infinite number of possible solutions may exist that satisfy the constraints. An objective or cost function is necessary to select the optimal configuration from among the set of feasible solutions. We use Beranek's ORM as our objective function, which is an application of Ando's Theory of Subjective Preference.[16,17] Ando found that when m orthogonal objective acoustic measures are given, the following definition of a cost function provides an acceptable scalar rating of a hall:

$$f(x) = \sum_{i=1}^{m} w_i f_i \tag{1}$$

Here, multidimensional vector x represents the configuration of the hall, function f_i penalizes the deviation from the target value of each objective acoustic measure, and weight w_i normalizes the respective functions.

Beranek uses the six objective acoustic measures (IACC, EDT, BR, G, TI, and SDI) and provides values for their weights, suitable for symphonic music.

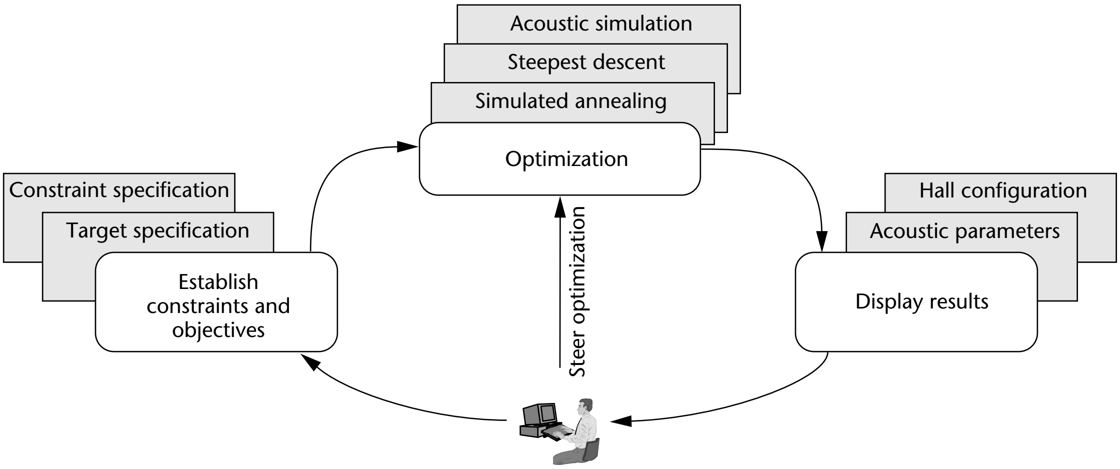
Finally, we minimize the objective function given by Equation 1 to find the hall configuration that best matches the target objective acoustic measures. Note that objectives may be constructed from all or a subset of the terms. In addition, it's possible to build objectives to handle multiple performance types simultaneously, such as symphonic music and opera. To accomplish this, we linearly combine individual objectives as follows:

$$f_{combined}(x) = a_1 f_{use_1}(x) + a_2 f_{use_2}(x) + \cdots a_n f_{use_n}(x)$$

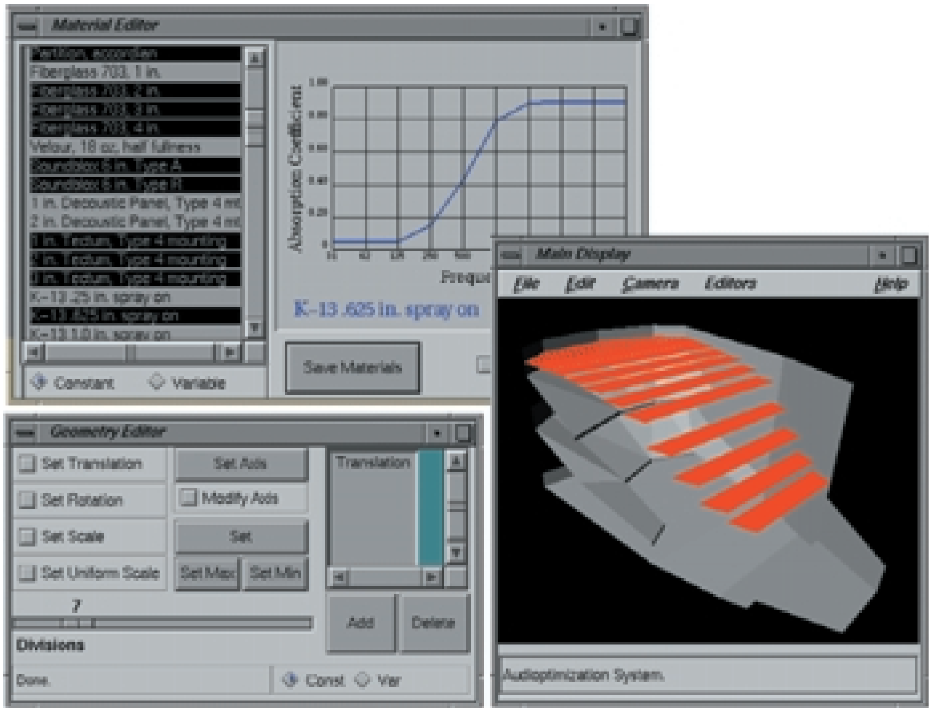
Here, a_i is the weighting factor of the ith individual objective function given n objectives, and the sum of the weighting factors is 1.0. Note that the minimum cost of a multiple-use objective function is typically not zero, since we cannot perfectly achieve all objectives.

Optimization problem

We state this problem as follows: minimize $f(x)$ subject to $x \in X$, where constraint set X is the "design space" spanned by feasible hall configurations. The existence of constraints implies that not every target is realizable. We



3 Overview of the interactive design process.



4 **Material and geometry editors.**

must identify an optimal point in the design space, that is, $x^* \in X$, such that $f(x^*) \leq f(x)$, $\forall x \in X$. We use simulated annealing and steepest descent techniques in combination to search globally for the "best" hall configuration.

Implementation

Figure 3 illustrates the framework of our approach in implementing an interactive design system based on audioptimization.

Users provide an initial model, which is passed to the simulation engine to compute a baseline sound field solution. Then, to specify desired targets for acoustic measures, they can manipulate glyphs and paint desired sound-level values onto a subset of hall surfaces at selected time steps. Users can then constrain the design space for the system to search, by indicating the range of modification for variable material and geometric components. They can optimize over materials and geometry either separately or simultaneously. Once the optimization process has been initiated, they can interrupt it to modify goals, add and/or delete variables and modify their constraints, or adjust optimization parameters. After all the design goals and variables are specified, the optimization process runs until convergence is achieved.

The following pseudocode describes the process:

```
Compute baseline sound field solution.
Establish constraints, objectives, and
   optimization parameters.
repeat
   Invoke simulated annealing.
   Invoke steepest descent.
   Display results.
   Modify constraints, objectives, and
      optimization parameters if
      desired.
until convergence.
```

Simulation

The sound field for the initial model configuration is simulated by the Monks hybrid simulation algorithm.[14] We use a standard source—a single omnidirectional, full-band spherical impulse—simulated and propagated into the environment. Users can trade simulation quality for speed through a number of interactive controls.

Visualizations described earlier display acoustic measures derived from the simulated sound field at a set of user-specified locations within the hall. A fine mesh of sample points displays sound-level data over time, as indicated by a color scale shown in Figure 2.

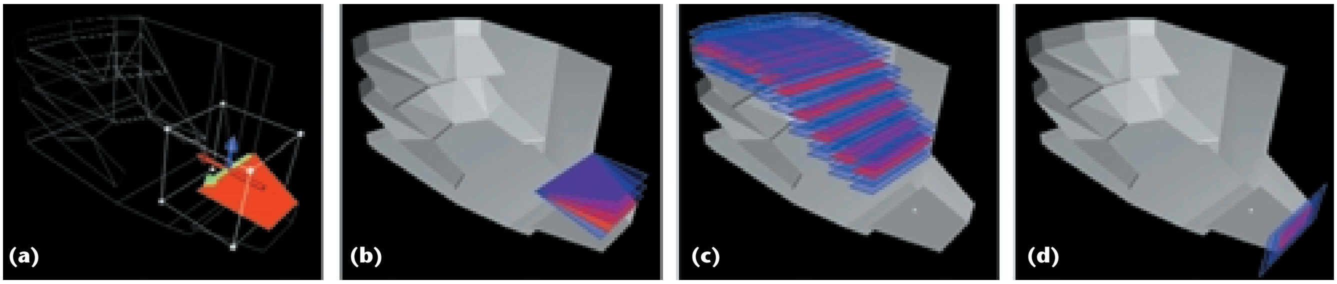
Constraints and objectives

Before initiating the optimization process, users specify a range of acceptable modifications to the hall. This specification involves selecting component modifications to surface materials and geometric transformations. Users must also set acoustic performance goals so the system can evaluate different hall configurations.

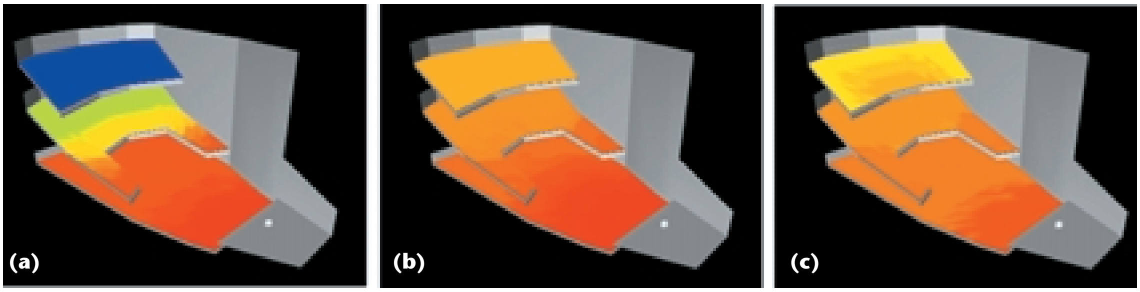
Constraint specification. To specify a material constraint for a component, users choose a set of allowable materials with the Material Editor (see Figure 4). An array of absorption coefficients corresponding to the frequencies, ranging from 31 Hz to 16 kHz, and an SDI value describe each material.

Users impose constraints on the transformation of geometric components using the Geometry Editor. After selecting a component and transformation type, they indicate the component's positional degree of freedom by placing and orienting the coordinate system axis icon. See Figure 5, next page. They set boundary constraints by positioning the component at the desired range limits. This range is discretized into a user-specified number of configurations.

The definition of a search neighborhood surrounding



5 Geometry constraint specification: (a) coordinate system axis icon used for transformation specification, (b) rotation constraint specified by orienting and selecting a rotation axis, (c) possible configurations resulting from translation constraint specification for a set of ceiling panels, and (d) scale constraint specified by indicating a point and direction of scale. The constraints are discretized according to user-specified divisions.



6 Sound strength target specification: (a) 80 ms, (b) 160 ms, and (c) total.

a geometry variable configuration comes naturally from this discretization process. The distance between configurations is simply the absolute difference between their configuration indices. The definition of a neighborhood surrounding a material variable configuration is less straightforward. The system orders materials based on their average absorption coefficients. The distance between configurations is then given by the absolute difference between their ordering indices. Because the ranges of materials and geometric transformations are discrete, combinatorial optimization is feasible, as we show later.

Target specification. To specify acoustic goals, designers directly manipulate the various sound field visualizations or enter scalar values. They set sound-level targets by selecting a paint color from the color palette and painting either a sound level or a change in level directly onto the surfaces for one or more time slices. Figure 6 illustrates desired sound levels for three time steps. To specify goals for BR, IACC, and EDT that vary by position, designers manipulate individual glyphs.

Optimization

In general, the audioptimization objective function is multiextremal, containing many local minima. Therefore, we first apply a global optimization step using simulated annealing to locate the neighborhood of a good solution, then follow with the steepest descent algorithm, which is more efficient in finding the local minimum.

Simulated annealing. Simulated annealing is a combinatorial optimization algorithm that produces a series of transitions between configurations in the design space based on three components: a generation mechanism, a cost function, and an acceptance function.[18] The generation mechanism randomly selects new configuration x_{new} from within a neighborhood around current configuration $x_{current}$, where the neighborhood size is determined by ε. Cost function f evaluates x_{new} as defined earlier. The acceptance function accepts or rejects x_{new} by comparing current cost $f(x_{current})$ to new cost $f(x_{new})$. If x_{new} is accepted, a transition results, and x_{new} replaces $x_{current}$. Unlike local optimization methods, which accept only lower cost transitions terminating at the local minimum, the simulated annealing method uses the Metropolis algorithm, in which the probability of accepting a higher cost configuration is nonzero.[19] This feature allows the search to proceed uphill, away from a local minimum, in search of a more global minimum.

Once the annealing process is initiated, it terminates upon satisfying one of two stop criteria. Either the process reaches a user-specified limit on the number of configurations to evaluate or the process converges. Convergence occurs after achieving an exhaustive evaluation and rejection of each configuration in the current neighborhood.

Once the stop criteria are met, users may assess the current state of the sound field. They may then conclude the optimization process by invoking the steepest descent step. Alternatively, they may alter acoustic performance goals, add or delete optimization variables or modify their constraints, and modify optimization parameters, then restart the optimization. If they have chosen not to make modifications and the optimization process has not yet converged, they may resume the optimization.

We express the simulated annealing algorithm in pseudocode as follows:

```
procedure SimulatedAnnealing()
   k ← 0
   AnnealingSchedule_InitParams(τk, εk, γk)
   xcurrent ← InitialRandomConfiguration()
   repeat
       for (i ← 0; i < γk; i ← i + 1)
           xnew ← GenerateNeighborConfiguration(εk, xcurrent)
           if (M (xnew, xcurrent, τk) > random[0,1))
               xcurrent ← xnew
           end if
       end for
       k ← k + 1
       AnnealingSchedule_UpdateParams(τk, εk, γk)
   until Stop Criteria
end procedure
```

Steepest descent. We follow simulated annealing with steepest descent. Starting at $x_{current}$, successive steps are taken between neighboring configurations with decreasing cost. The algorithm terminates when no neighboring configuration has a lower cost than the current configuration.

We determine descent direction by computing n-dimensional direction vector d, where n is the number of optimization variables. Each vector component d_i, where $i = 0, \ldots, n-1$, may assume the value −1, 0, or 1, depending on which neighboring configuration has the lowest cost. We locate the most suitable neighbor by defining unit direction vector e, setting its ith component to 1, adding e to and subtracting e from $x_{current}$ in turn, and comparing the cost of the resulting configurations. The pseudocode in Figure 7 shows this discrete version of the steepest descent algorithm.

```
procedure SteepestDescent()
  repeat
    d ← 0
    for (i ← 0; i < n; i ← i + 1)
        e ← 0
        ei ← 1
        if ((f (xcurrent − e) < f (current) and
          (f (xcurrent − e) < f (xcurrent + e)))
           di ← −1
        else if ((f(xcurrent + e)<f(x current t)) and
          (f (xcurrent + e) < f (current − e)))
           di ← 1
        else
           di ← 0
        end if
    end for
    if f (xcurrent + d) < f (current)
        xcurrent ← xcurrent + d
    end if
  until xcurrent is optimal
end procedure
```

7 **Pseudocode of the discrete version of the steepest descent algorithm.**

Discussion. When optimizing over both materials and geometry, we can alter the optimization algorithm somewhat. We can take advantage of the fact that material modifications execute up to 50 times faster than geometry modifications. Therefore, we run a full material optimization after testing a new geometry configuration and use the resulting cost in the acceptance test for the geometry optimization. The second example in the "Case studies" section uses this approach.

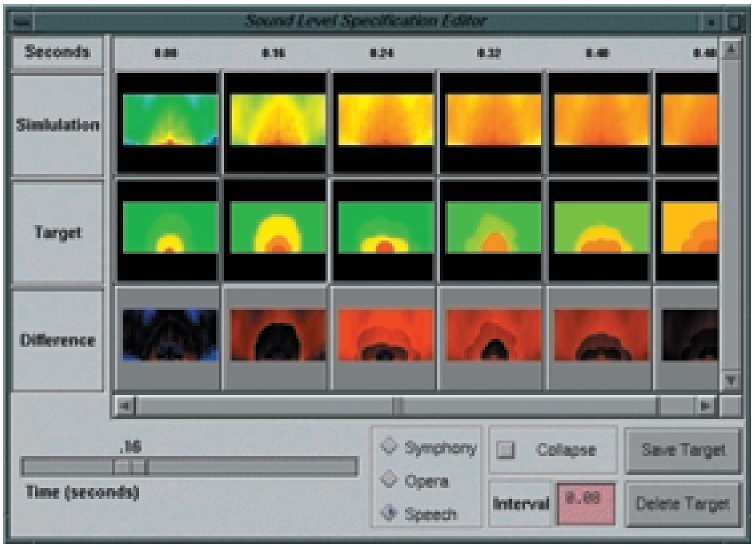
Simulated annealing has both advantages and disadvantages. On the positive side it has a statistical guarantee of locating an optimal solution, and the objective function may be discontinuous and nondifferentiable. This approach is well suited to searching discretized design spaces such as ours.

On the negative side the control parameters need tuning for each new application to obtain the best results. In our case the cooling schedule required to guarantee statistically locating an optimal solution is impractically slow. As with many real-life applications of simulated annealing, we relax this condition, instead using simulated quenching.[20] It provides an acceptable solution within a reasonable time frame.

Display results

After the optimization process modifies the hall, we assess the resulting acoustics with a variety of visualization tools, which convey the current state of the sound field. We have augmented these displays with visualizations that show the difference between the current sound field and the objectives.

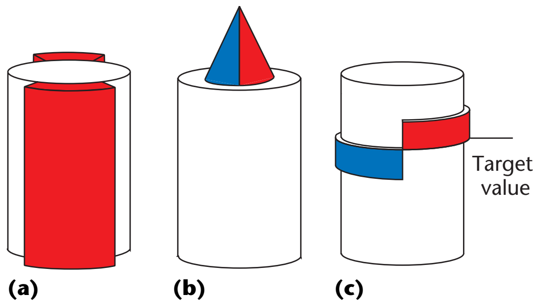
A scrollable viewer indicates sound level with a column of three thumbnail images for each time slice (Figure 8). The top row shows the color-coded visualization



8 **Sound-level specification editor: (top row) simulated sound strength data, (middle row) painted targets, and (bottom row) signed differences for a set of time steps.**

for the actual sound level. The middle row shows target values specified for a given time slice. At the bottom is a red-blue signed difference image between the actual and target values, where difference is defined as the actual value minus the target value.

Users define the mapping by specifying the range limits represented. The maximum positive difference is mapped to red, maximum negative difference to blue, and minimum difference to black. This tool provides an easy way to assess and specify desired performance through time.



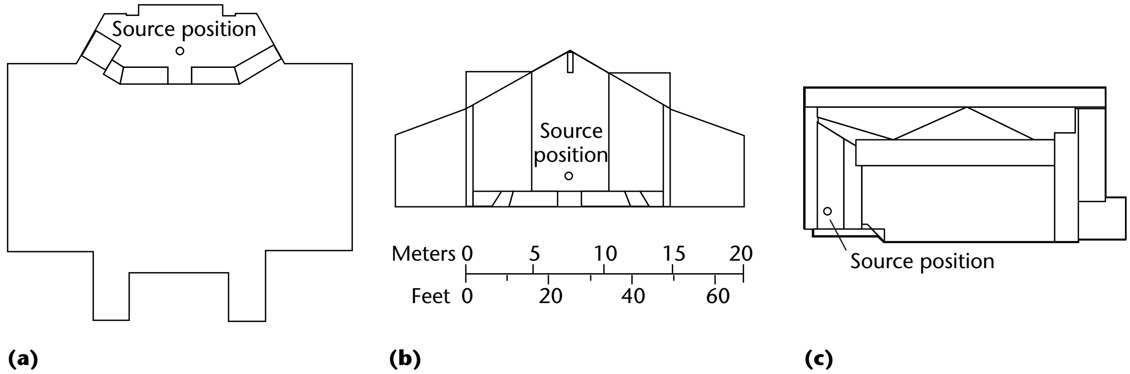
9 Graphical difference glyphs representing (a) IACC, (b) EDT, and (c) BR.

Figure 9 shows positional difference glyphs. We represent IACC difference as the region not covered by the actual value. Under ideal conditions the IACC difference shell is absent, since it represents directions *not* covered by the incoming sound. We characterize EDT difference as a solid cone whose radius indicates absolute difference and whose color indicates whether the actual value is higher than (red) or lower than (blue) its target. We represent BR difference as a ring between the actual and target values, similarly color-coded.

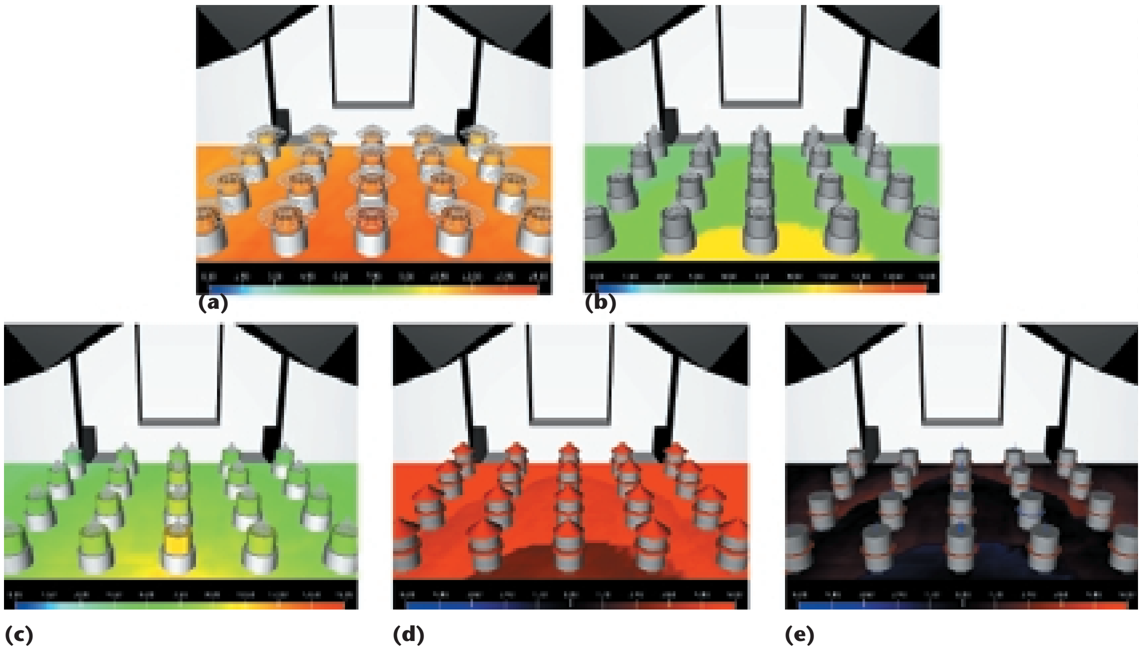
Case studies

We implemented our system in C++ and ran it on an SGI Onyx RealityEngine2 workstation with 256 Mbytes of memory using a single 195-MHz MIPS R10000 processor.

To demonstrate the inverse approach, we show results based on two real architectural spaces in which acoustic considerations play a prominent role. Both examples involve a common acoustic design process—the renovation of a flawed structure to correct acoustic problems.



10 Computer model of Oakridge Chapel: (a) plan, (b) cross section, and (c) longitudinal section.



11 EDT, BR, and G values for initial, target, and optimized configurations of Oakridge Bible Chapel: (a) initial simulation, (b) target, (c) final simulation, (d) initial difference, and (e) final difference. Material variables include the walls and ceiling.

Oakridge Bible Chapel

Oakridge Bible Chapel in Toronto, Canada, exhibited a fundamental acoustic flaw in its initial design (see Figure 10). The walls and ceiling were faced with highly reflective plaster, which produced excessive reverberation. This caused the sound from one spoken syllable to linger and mask the following syllable. Even when sufficiently loud, speech became almost unintelligible.

We ran a simulation of the existing environment, which confirmed the speech intelligibility problem. Figures 11 through 13 show the results of the simulation. The width of the cones and height of the bass cylinders indicate that both EDT and BR are too high (see Figure 11a). The sound strength visualization shown in Figure 13 (next page) demonstrates that the total sound level is also too high and that much of the distribution arrives late. The time distribution of sound energy also has a significant effect on speech intelligibility; the earlier the sound arrives, the better.

We set out to improve speech intelligibility by reducing the initially high values of EDT and BR while maintaining adequate sound levels. We built our objective function using EDT, BR, and G, which are the three measures relevant for speech. Table 1 lists their respective target values. Note that these values differ from those used to evaluate symphonic music.[1] Figure 11b shows the target values, and Figure 11d shows a difference image relative to the initial configuration. To improve speech intelligibility, we painted sound-level target distributions on the seating plane for three time slices: 0.08 seconds, 0.16 seconds, and total level, as shown in Figure 12. We chose these times slices for our target to include reflected sound.

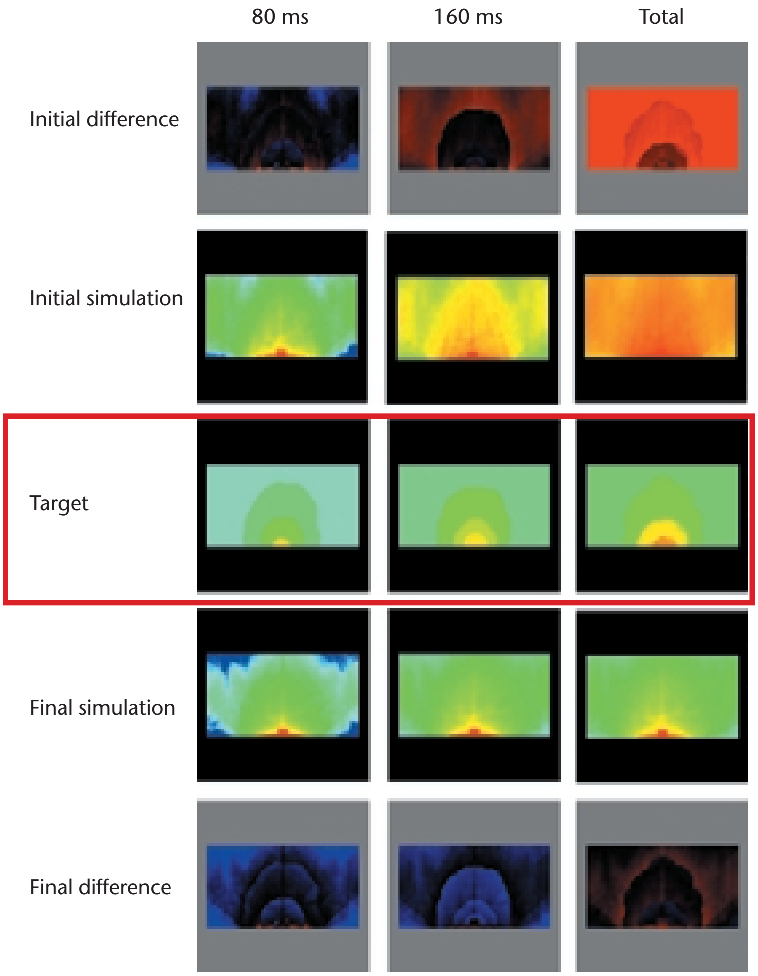
Having set our acoustic goals, we selected modifiable components and specified the range of the modifications. Oakridge is typical of buildings constrained by their existing geometry, limiting redesign options. With this in mind, we restricted our optimization to include only changes to materials. We considered two design scenarios: one involving the ceiling surfaces—the most easily modifiable surfaces covering the largest contiguous area—and the second involving both the ceiling surfaces and walls.

In the first scenario, the system assigned highly absorptive materials to cover the reflective ceiling surfaces. The resulting configuration yielded an objective value 97.5 percent closer to our goal than the initial configuration. As desired, EDT dropped greatly to a fraction of the original value, and BR dropped significantly as well. Speech would be much more intelligible to the congregation with the resulting configuration.

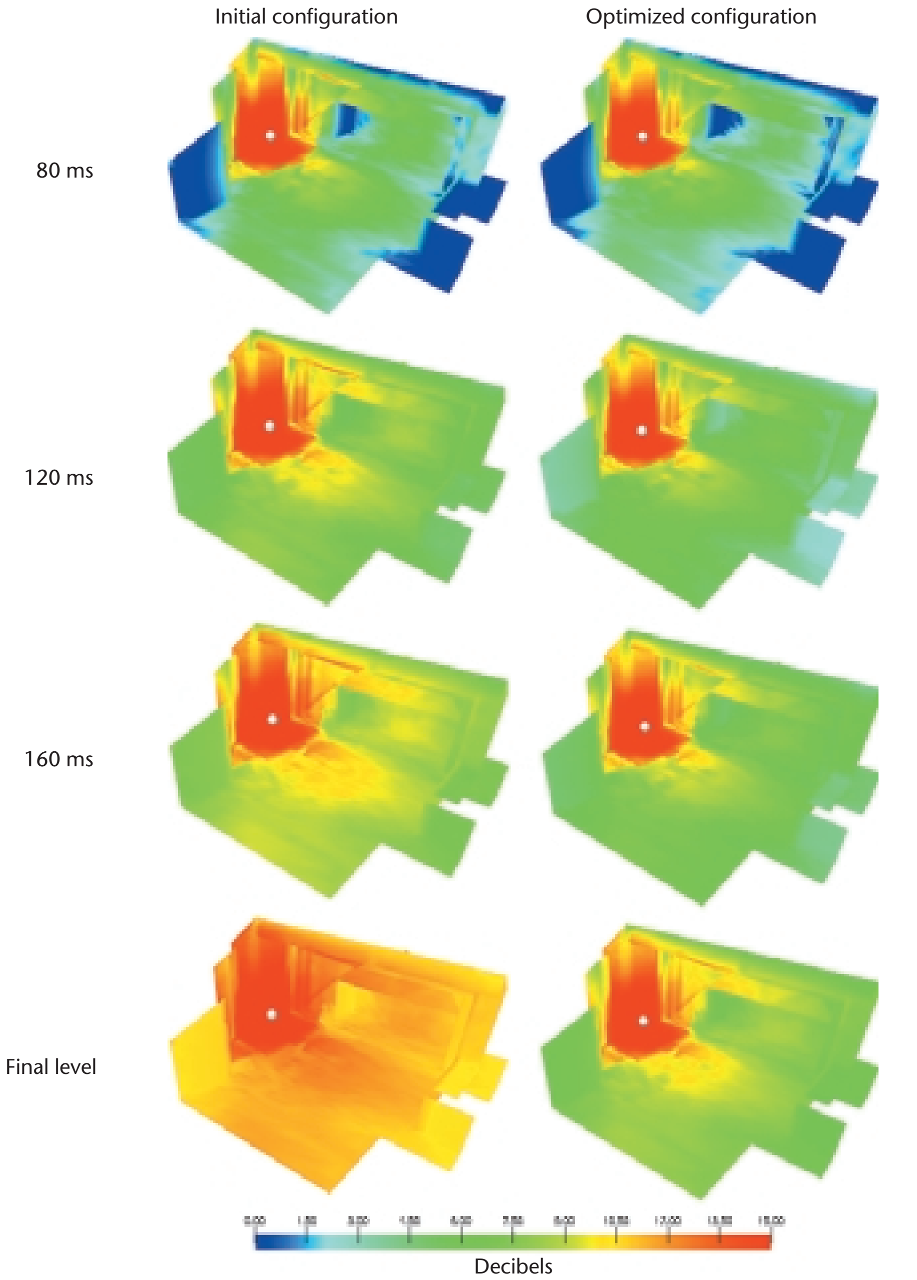
For the second scenario, we also altered the materials on the walls. The system assigned highly absorptive materials to most surfaces, which improved the objective value by 98.2 percent. Figure 13 shows the accumulation of sound energy on the seating plane at four time steps for both the initial and final configurations. Note that in the final configuration, the majority of the sound arrives early, resulting in improved speech intelligibility. We have in effect improved the temporal distribution of sound by increasing the percentage of the early sound and decreasing the percentage of late sound. Table 1 summarizes the results. Each scenario ran in under 3 minutes.

Table 1. Acoustic measure readings for Oakridge Chapel.

	EDT(s)	BR	G(dB)	$f(x)$	Improvement (%)
Target	0.700	1.000	≥ 0.0	0.0	N/A
Initial Configuration	1.960	1.617	12.103	5.140	N/A
Final (ceiling only)	0.787	1.007	7.458	0.130	97.5
Final (walls and ceiling)	0.646	1.034	6.652	0.093	98.2



12 Sound strength shown in the seating areas of Oakridge Bible Chapel at three time steps for initial, target, and optimized configurations. Material variables include the walls and ceiling.



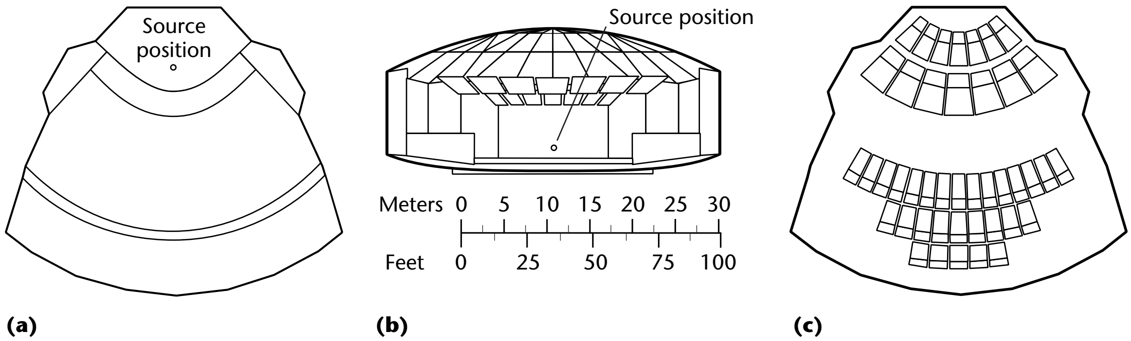
13 Sound strength shown interior surfaces of Oakridge Bible Chapel at four time steps for initial and optimized configurations. Material variables include the walls and ceiling.

Kresge Auditorium

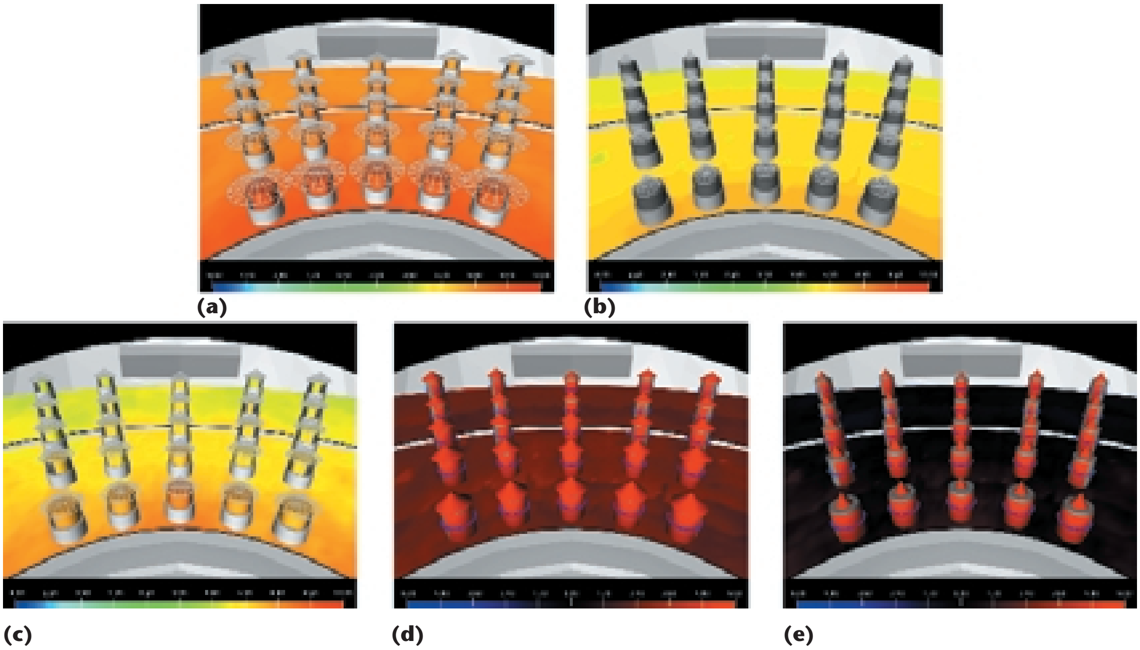
The second example, Kresge Auditorium, is a multipurpose auditorium at MIT, which is currently undergoing acoustic reevaluation (see Figure 14). The Institute uses Kresge Auditorium for everything from conferences to concerts. The hall doesn't possess reconfigurable elements that would help accommodate such disparate acoustic requirements. Consequently, the auditorium suffers from too much reverberation for speech, although the reverberation is adequate for music, as shown graphically in Figure 15 by the EDT cones. Another shortcoming is that the audience doesn't feel surrounded by the sound, indicated by the IACC shells, which fail to encircle the glyph. The average sound level G is 7.33 dB (see Tables 2-4). However, the temporal distribution of energy is poor for speech intelligibility, with too much energy arriving late.

While the hall's acoustics will never satisfy all uses without introducing reconfigurable geometry or material elements, our intent was to consider modifications that would improve the acoustics as a whole. To reflect this, we modified our objective function to include two sets of targets, one each for speech and symphonic music, equally weighted. We painted sound-level targets for both speech and symphonic music for time slices at 0.08 seconds, 0.16 seconds, and total level, as shown in Figure 16 (page 88). Table 2 gives the other targets.

Unlike the Oakridge Chapel optimization, which we restricted exclusively to material changes, here we include optimization over geometry



14 Computer model of Kresge Auditorium: (a) plan, (b) cross section, and (c) reflected ceiling plan.



15 **IACC, EDT, BR, and G values for initial, target, and optimized configurations of Kresge Auditorium using the combined objective for speech and symphonic music: (a) initial simulation, (b) target, (c) final simulation, (d) initial difference, and (e) final difference. Material variables include the seats, stage walls, and far wall. Geometry variables include the rotation of the stage ceiling reflectors and the translation of the rear stage well.**

Table 2. Acoustic measure readings for Kresge Auditorium: speech and symphonic music.

	IACC	EDT (s)	BR	G (dB)	SDI	TI (s)	*f*(*x*)	Improvement (%)
Target	0.000	0.850	1.000	4.750	1.000	0.020	0.830	N/A
Initial Configuration	0.618	2.178	0.939	7.325	0.202	0.039	3.302	N/A
Final (materials only)	0.721	1.401	1.049	3.102	0.219	0.039	2.276	41.5
Final (geometry only)	0.537	2.098	0.945	7.364	0.202	0.028	3.054	10.0
Final (both)	0.592	1.479	0.973	4.989	0.198	0.028	2.234	47.8

Table 3. Acoustic measure readings for Kresge Auditorium: speech.

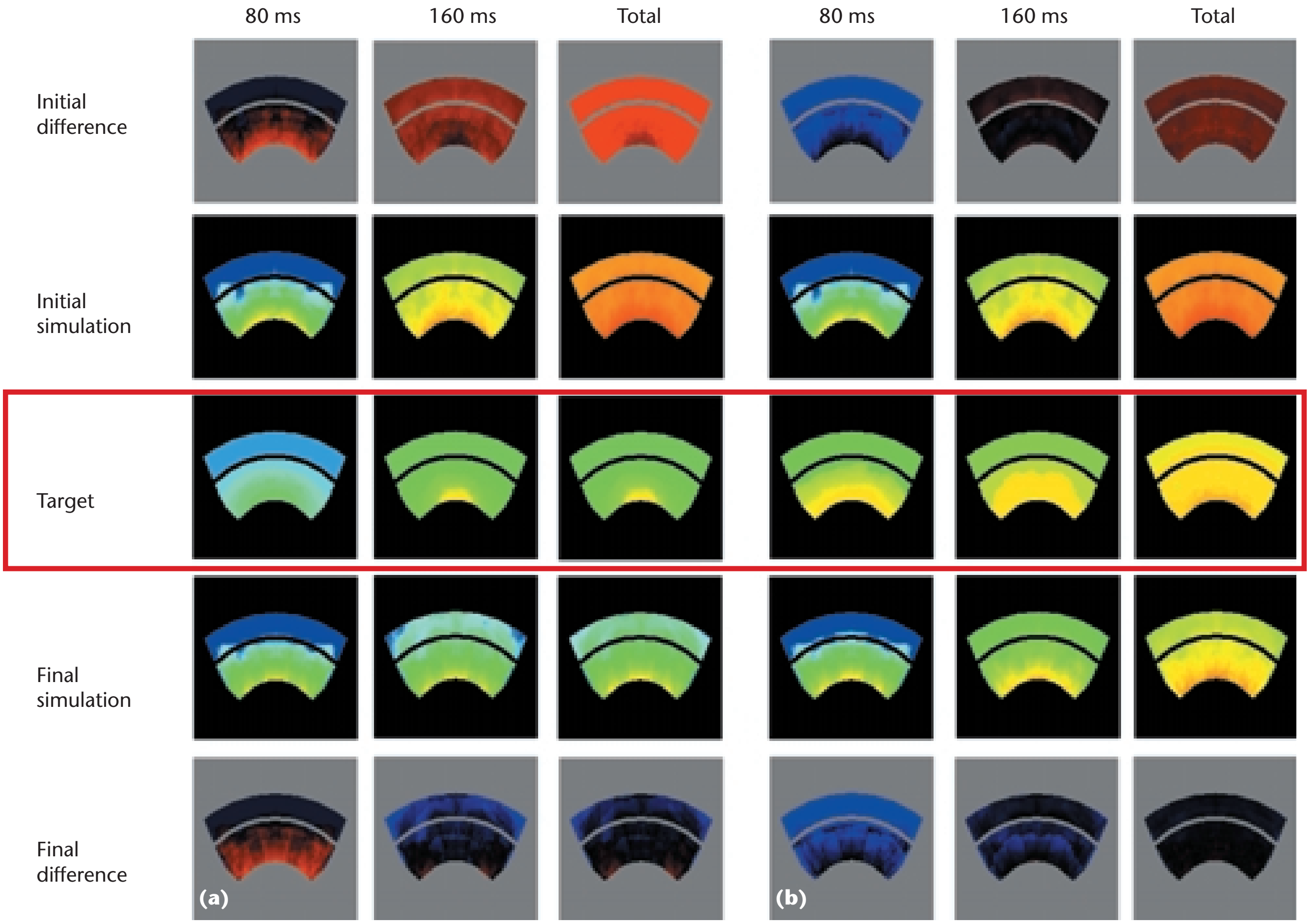
	IACC	EDT (s)	BR	G (dB)	SDI	TI (s)	*f*(*x*)	Improvement (%)
Target	N/A	0.700	1.000	≥0.000	N/A	N/A	0.0	N/A
Initial Configuration	N/A	2.178	0.939	7.325	N/A	N/A	4.621	N/A
Final (both)	N/A	0.785	1.025	0.150	N/A	N/A	0.296	93.6

Table 4. Acoustic measure readings for Kresge Auditorium: symphonic music.

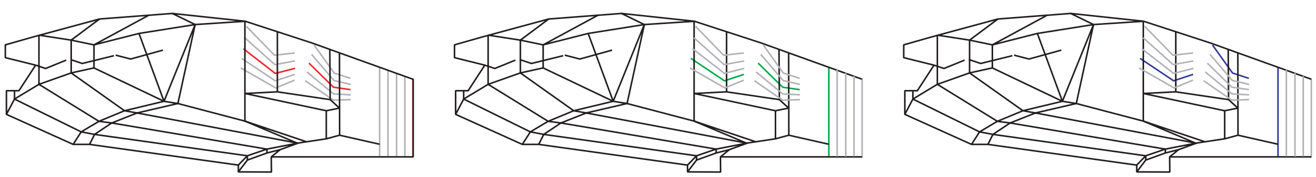
	IACC	EDT (s)	BR	G (dB)	SDI	TI (s)	*f*(*x*)	Improvement (%)
Target	0.000	2.150	1.175	4.750	1.000	0.020	0.0	N/A
Initial Configuration	0.618	2.178	0.939	7.325	0.202	0.039	1.983	N/A
Final (both)	0.588	1.947	1.013	4.485	0.293	0.028	1.350	31.9

as well. Since construction costs for geometric changes generally exceed costs for material changes, we separated the optimization into three passes—modifying only materials, only geometry, then materials and geometry combined—to compare the effectiveness of each.

Using the visualization tools, we observed the pattern of sound-level accumulation as a function of time for the initial hall configuration. We noted that the direct sound and the earliest reflections have the greatest effect. As variables for the first example, we selected the



16 **Sound strength shown on the seating areas of Kresge Auditorium at three time steps for initial and optimized configurations using separate (a) speech and (b) symphonic music objectives. Material variables include the rotation of the stage ceiling reflectors and the translation of the rear stage wall.**

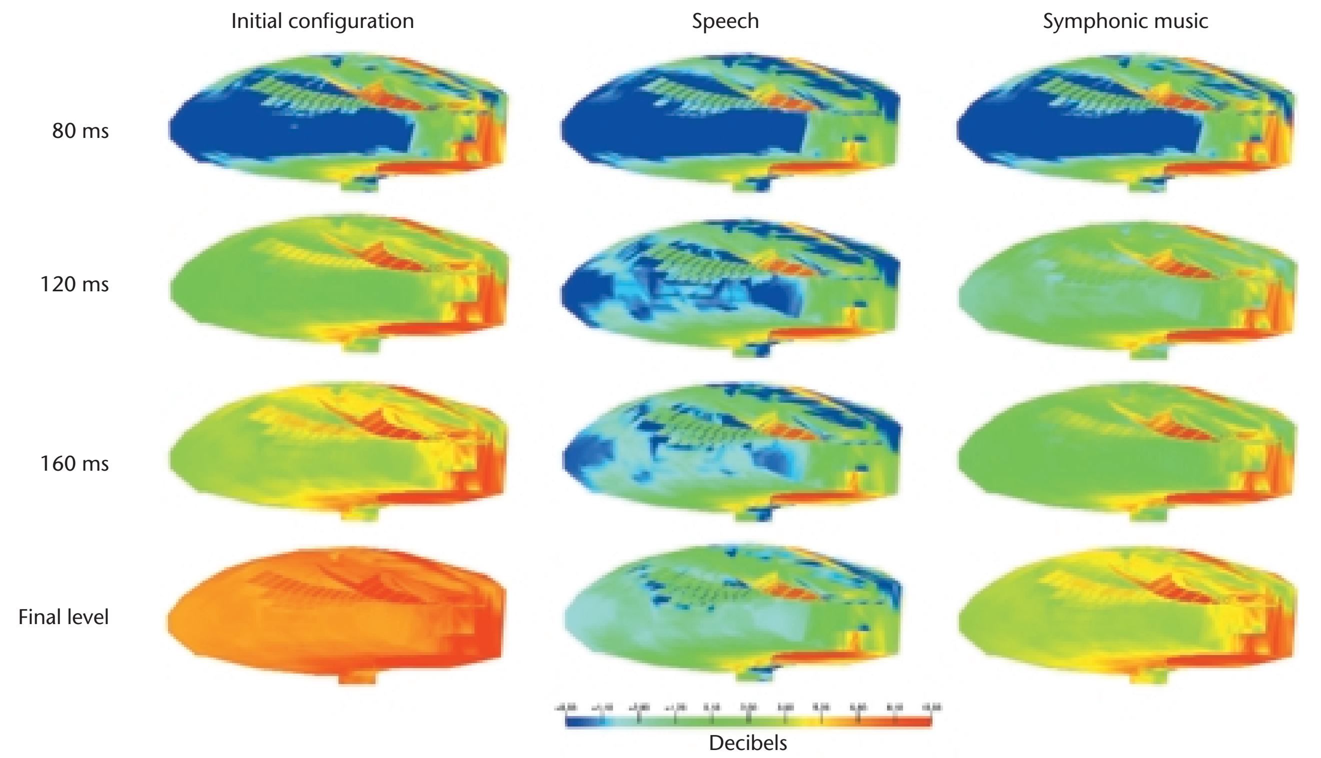


17 **Variable positions for the rear and forward bank of reflectors and the back stage wall in Kresge Auditorium for (a) the initial (red) configuration, (b) the geometry only (green) configuration, and (c) the combined materials and geometry (blue) configuration.**

materials for the seats, the wall at the back of the hall, and the walls of the stage shell, since reflected sound from these surfaces reaches much of the seating area.

The optimization over materials took 72 seconds to converge, sampling 200 configurations. The system assigned absorptive material to the seats and stage floor, and reflective materials to the rear wall and remaining surfaces near the stage. The Materials Only entry in Table 2 shows that EDT improved substantially, and BR nearly matched the target. While SDI improved, we did not include surface treatments that would raise SDI much beyond 0.3. TI remained unchanged, since it is only affected by geometric changes.

In the second example, the geometry modifications included the depth of the center stage wall and the rotation of the two sets of suspended reflector groups above the stage area, as illustrated in Figure 17. Each component could assume one of five positions, with the initial configuration indicated in red. These geometric components share the characteristic that the ratio between their size and the solid angle they span with respect to the sound source location is small. Further, these mod-



18 **Sound strength shown on the interior surfaces of Kresge Auditorium at four time steps for initial and optimized configurations using speech and symphonic music objectives separately. Material variables include the seats, stage walls, and far wall. Geometry variables include the rotation of the stage ceiling reflectors and the translation of the rear stage wall.**

ifications wouldn't require expensive alterations to the external shell of the auditorium. In the optimization over geometry, the system left the orientation of the rear reflector group unchanged, but lowered the forward reflector group and moved the rear stage wall to the position closest to the source, as shown in green in Figure 17b. These modifications improved TI by 58 percent. The optimization took approximately 20 minutes to converge while sampling 100 geometric configurations.

The combined optimization—involving both materials and geometry—altered materials as before. However, it selected a new configuration for the banks of reflectors, raising the rear reflector group, lowering the forward reflector group, and again moving the rear stage wall to the position closest to the source, as shown in blue in Figure 17c. This configuration produced the lowest cost by maintaining the improvements to TI from geometry modifications and improvements to EDT from material modifications. Sound-level G dropped to 5.0 dB and the temporal distribution improved, with a higher percentage of the energy arriving earlier than in the initial configuration. IACC improved somewhat, but remained far from optimal, which is expected for fan-shaped halls such as this one. The optimization took 17 minutes to converge while sampling 80 geometric configurations, improving the overall acoustic rating by 47.8 percent. The results in Table 2 and Figure 15 show that performance improved for both uses.

Finally, taking the resulting geometric configuration from the combined optimization, we introduced a set of materials that could be changed between speech and symphonic music performances. Examples are curtains that could be drawn or rugs that could be taken up. We restricted our study to stage surfaces and the auditorium's back wall. The system assigned highly absorptive materials for the speech configuration and materials with mid to low absorption for the music configuration. Tables 3 and 4 and Figures 16 and 18 show the improvements for speech and music respectively at 93.6 percent and at 31.9 percent.

Our interactive system for acoustic design solves a restricted inverse problem. The user provides a specification of desired acoustic performance and describes material and geometric variations and constraints for a collection of architectural components in the scene. The system then searches the design space for the configuration that "best" meets the specification. This approach can be easier and more intuitive to use than the usual direct edit-simulate cycle.

Our experiences with the system suggest several areas for future work:

- *Application to preliminary design.* While we have tested the system on models of existing halls, we are eager to explore the system's utility as a tool in the design of a new hall. Because the acoustic design problem is so difficult, designers often feel compelled to use a familiar or proven geometric configuration to avoid a potentially costly mistake. Our system seems amenable to assisting in the preliminary design phase—perhaps enabling a designer to not only consider a wide range of possible designs but also to gain

One intriguing area for future work is in visualizing the optimization process itself—providing the designer with an intuitive representation of the multidimensional search space and the ability to steer the optimization process.

new insight into the intricacies of sound propagation.

- *Acoustic simulation and visualization.* Currently, the sound field is sampled at a uniform grid of points, each contributing equally to the evaluation of the hall. There may be better ways of combining the sample data, perhaps giving more weight to central seating area locations. In addition, the area of accurate acoustic simulation is an important one, with a variety of challenges remaining such as diffraction and other wave effects. Volumetric representations of the sound field might also improve the user's understanding of the acoustics but would require a voxel representation of the 3D scene.
- *Auralization.* Some systems play back a prerecorded anechoic sound sample convolved with the sound field signature at a specific location within the hall, permitting the listener to hear what a hall might sound like at that location. We could use this approach and extend it in our system. The listener could set acoustic targets by modifying the sound-field signature interactively while listening to the resulting signal, using graphical tools (sliders or dials hooked to specific sound field characteristics). While one can only listen to the sound produced at a single location in the hall at a time, targets can influence any user-specified region.
- *Optimization and design.* The advent of computer-aided design systems has brought designers tools that assist in predicting and visualizing complex phenomena such as light and sound. Clearly, however, the computer could play a more significant role in the design process. One intriguing area for future work is in visualizing the optimization process itself—providing the designer with an intuitive representation of the multidimensional search space and the ability to steer the optimization process. Perhaps the greatest shortcoming of the current system is that it cannot directly identify the components that have the strongest adverse effect on acoustic performance. Such variable sensitivity analysis, in addition to boundary analysis and application-specific heuristics based on the insights of experienced acoustical engineers, might prove beneficial. Finally, the Design Galleries[5] approach and our optimization approach might be combined. The dispersion phase of DG could select configurations that not only meet a difference criteria but also satisfy evaluation requirements.

Although they have received little attention in computer graphics to date, inverse algorithms have great potential as design tools. ■

Acknowledgments

An Alfred P. Sloan Research Fellowship (BR-3659) and a grant from the National Science Foundation (IRI-9596074) supported this work. Kurt Wagner of the Bose Corporation provided the Oakridge Bible Chapel model. We thank Steven Gortler and Carl Rosenberg for helpful discussions and Max Chen for experimenting with the system and modeling Kresge Auditorium. We especially thank the reviewers of this document for their many helpful comments and suggestions.

References

1. M. Barron, *Auditorium Acoustics and Architectural Design*, E. and F.N. Spon, London, 1993.
2. J. Borish, "Extension of the Image Model to Arbitrary Polyhedra," *J. Acoustic Society of America*, Vol. 75, 1984, pp. 1827-1836.
3. T. Lewers, "A Combined Beam Tracing and Radiant Exchange Computer Model for Room Acoustics," *Applied Acoustics,* Vol. 38, 1993, pp. 161-178.
4. Bose Corporation, Auditioner Audio Demonstrator Technology—In-Depth, Nov. 1998, http://www.bose.com/technologies/prediction_simulation/html/principle.html.
5. J. Marks et al., "Design Galleries: A General Approach to Setting Parameters for Computer Graphics and Animation," *Computer Graphics Proc., Ann. Conf. Series*, ACM Press, New York, Aug. 1997, pp. 389-400.
6. P. Poulin and A. Fournier, "Lights From Highlights and Shadows," *Proc. 1992 Symp. Interactive 3D Graphics*, D. Zeltzer, ed., Apr. 1992, ISBN 0-89791-467-8, pp. 31-38.
7. C. Schoeneman et al., "Painting With Light," *Computer Graphics Proc., Ann. Conf. Series*, ACM Press, New York, 1993, pp. 143-146.
8. J.K. Kawai et al., "Radioptimization—Goal-Based Rendering, *Computer Graphics Proc., Ann. Conf. Series*, ACM Press, New York, 1993, pp. 147-154.
9. D.P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, Massachusetts, 1995.
10. N. Tsingos, *Simulating High Quality Virtual Sound Fields for Interactive Graphics Applications,* PhD thesis, Comp. Sci. Dept., Univ. J. Fourier, Grenoble, France, Dec. 1998.
11. A. Krokstad et al., "Calculating the Acoustical Room Response by the Use of a Ray Tracing Technique," *J. Sound Vibration,* Vol. 8, No. 1, 1968, pp. 118-125.
12. P.S. Heckbert and P. Hanrahan, "Beam Tracing Polygonal Objects," Computer *Graphics* (Proc. Siggraph 84), H. Christiansen, ed., Vol. 18, ACM Press, July 1984, pp. 119-127.
13. T. Funkhouser et al., "A Beam Tracing Approach to Acoustic Modeling for Interactive Virtual Environments," *Computer Graphics Proc., Ann. Conf. Series*, ACM Press, New York, July 1998, pp. 21-32.
14. M. Monks, "Acoustic Simulation and Visualization Using a New Unified Beam Tracing and Image Source Approach," *Proc. Convention Audio Engineering Society*, (preprint), 1996; copies available from the authors.
15. A. Stettner and D.P. Greenberg, "Computer Graphics Visu-

alization for Acoustic Simulation," *Computer Graphics* (Proc. Siggraph 89), J. Lane, ed., ACM Press, New York, Vol. 23, July 1989, pp. 195-206.
16. L.L. Beranek, *Concert and Opera Halls: How They Sound*, Acoustical Society of America, New York, 1996.
17. Y. Ando, *Concert Hall Acoustics,* Springer-Verlag, Berlin, 1985.
18. S. Kirkpatrick et al., "Optimization by Simulated Annealing," *Science,* Vol. 220, 1983, pp. 671-680.
19. E. Aarts and J. Korst, *Simulated Annealing and Boltzman Machines*, John Wiley & Sons, New York, 1989.
20. L. Ingber, "Simulated Annealing: Practice Versus Theory," *Mathematical Computing Modeling,* Vol. 18, 1993, pp. 29-57.

Michael Monks *is a senior software engineer at Bose Corporation. His interests include acoustic system design, as well as parametric and procedural modeling. He received a PhD in computer graphics from the Massachusetts Institute of Technology, an MS in computer graphics from Cornell University, and a BS in computer and information science from the University of Massachusetts, Amherst.*

Byong Mok Oh *is a PhD student in the Computer Graphics Group at the Massachusetts Institute of Technology. His research interests include visualization and rendering techniques. He received his MSE in computer information sciences from the University of Pennsylvania and BA degrees in computer science and in studio art and art history from Oberlin College.*

Julie Dorsey *is an associate professor in the Departments of Electrical Engineering and Computer Science and Architecture and a member of the Laboratory for Computer Science at the Massachusetts Institute of Technology. Her research interests include rendering algorithms, material and texture models, illustration techniques, and interactive visualization of large databases, with an application to urban environments. She received BS, BArch, MS, and PhD degrees from Cornell University. In 1996, she was awarded a Faculty Early Career Development Award from the National Science Foundation, and in 1997 an Alfred P. Sloan Foundation Research Fellowship and the Harold E. Edgerton Faculty Achievement Award from MIT. She is a member of the ACM.*

Readers may contact Michael Monks at Bose Corporation, The Mountain, Framingham, MA 01701-9168, e-mail michael_monks@bose.com.