Designing an Urban-Scale Auditory Alert System

Drawing on a wide range of computing technologies and methodologies, the authors present a new auditory alert system for high tides in Venice designed to replace the existing network of electromechanical sirens.

Federico Avanzini
University of Padova

Davide Rocchesso

Alberto Belussi
University of Verona

Alessandro Dal Palù

Agostino Dovier
University of Udine

The high tides that periodically flood Venice, locally known as *acqua alta,* are becoming more serious due to recent changes in the surrounding lagoon as well as atmospheric conditions. Several days each year, tidal waters cover many city streets and squares, disrupting the inhabitants' lives and snarling traffic. Major floods, like the ones that struck Venice in 1966 and 2000, can inundate the city and paralyze activity.

Given the severity and increasing frequency of tidal flooding, alerting the population promptly has become imperative so that public officials, merchants, and citizens can take appropriate measures. A special office of the Municipality of Venice, the Center for Tide Prediction and Warning (Centro Previsioni e Segnalazioni Maree—CPSM), provides a continuous tide forecast based on computational models as well as astronomical and meteorological data. When a significant high tide is expected, city authorities activate a network of electromechanical sirens for a few minutes, usually anticipating the tide peak by a few hours.

The sirens, however, emit threatening wails reminiscent of air attack warnings, do not convey the gravity of the threat, and may not reach isolated or distant areas. For these reasons, the CPSM, in cooperation with the Venetian Research Consortium and the University of Verona, is investigating the possibility of replacing the sirens with a loudspeaker system that would provide more uniform coverage as well as information about the tide level. Although loudspeakers are more noticeable than sirens—a nontrivial concern in a city celebrated for its visual charms—they make it possible to broadcast any kind of sound.

As part of this research effort, our project team first analyzed the current alert system using off-the-shelf acoustic simulation software and a specially designed visualization tool. We then used a form of constraint logic programming to determine the optimal placement of loudspeakers in Venice, a complex task with many physical, economic, and social constraints. Next, we created the alert sounds for our demanding listening environment. The final phase of the project involved iteratively validating and redesigning the alert signals using human testing.

ACOUSTIC SIMULATION AND VISUALIZATION

The initial phase of our study consisted of a detailed analysis of the existing alert system. We first developed a technique that semiautomatically extracts building and terrain data from digital city maps in ArcView format with reasonable confidence. Our extraction technique structures this data as polygons representing land, water, and buildings with their associated height. We then imported this data into SoundPLAN (www.soundplan.com), an integrated software package for noise and air pollution simulations, and used it to generate a map of the sirens' sound-pressure levels throughout Venice.

We modeled the city's eight electromechanical sirens as point sources with an omnidirectional propagation pattern and used Fourier analysis of a steady-state portion of a siren sound to determine

Figure 1. Noise map of alert system currently used in Venice. Colors represent sound-pressure level in dB(A). The sirens' coverage is far from uniform, and in many areas the sound levels are inadequate.

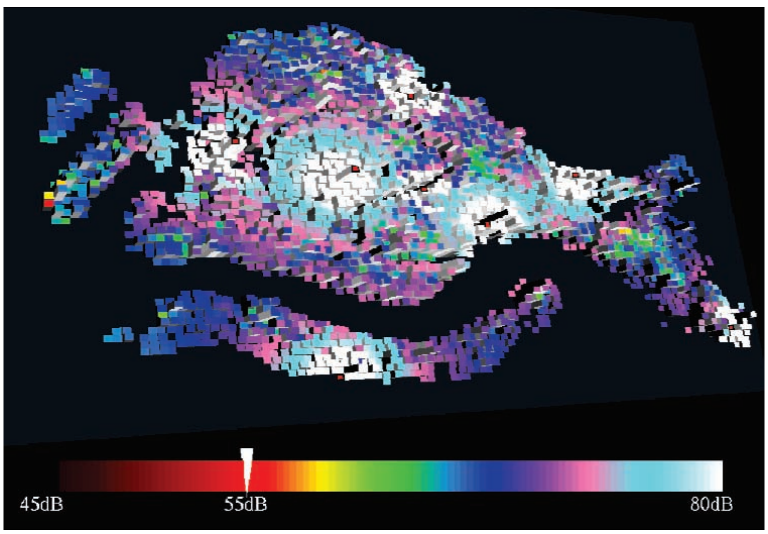


Figure 2. Simulation of alert system with simplified acoustic description. Compared with the noise map shown in Figure 1, the simplified description provides slightly overestimated values (in dB) in a very short computation time.

their spectral content. As Figure 1 shows, the sirens' coverage is far from uniform, and in many areas the sound levels are inadequate.

SoundPLAN includes a ray-tracing algorithm that computes acoustic effects such as reflection, refraction, absorption, and shielding. As a tradeoff between accuracy and feasibility, our simulations used a 5-meter grid step, a value larger than many Venetian channels and alleys. Even with such a large step and a reasonable number of ray reflections, four, the algorithm takes days to compute. Thus, while SoundPLAN is useful for producing a reliable image of a given situation and assessing the validity of proposed solutions, using it as an exploratory tool or routinely embedding it in optimization procedures is inappropriate.

We therefore developed a simplified acoustic description of the current alert system, in the form of a multilayer grid, that assumes the phases of signals emanating from different sources randomly mix at the listening point. This assumption—especially valid in complex urban environments—allows separate computation of each source's sound field regardless of the nature of the signals the system is emitting, and constructive summation of component tones' intensities.[1]

Given a source located in cell s of the grid, the model assumes that the power level L_s dB of s is known. The environmental attenuation in the jth cell is computed by calculating individual attenuation components $K_i(d_{s,j})$ dB along the distance $d_{s,j}$ between cells s and j. The resulting sound-pressure level at point p is then computed using the discrete function

$$f_s(j) = L_s - \sum_i K_i\left(d_{s,j}\right) \text{ dB}.$$

Many noise-prediction schemes take this modeling approach. One of the most influential implementations is the German Association of Engineers' VDI 2714/2720 standard (www.vdi.de). The function f_s exploits the specifications of this standard to compute the K_i components that account for

- attenuation due to free-field propagation,
- additional air absorption, and
- shielding due to buildings.

To validate our simplified acoustic description of the current alert system, we performed a simulation of the existing setup. We used the specially written OpenGL-based application shown in Figure 2 to visualize the simulation and compared the results with those obtained via SoundPLAN. Although the units in Figures 1 and 2 differ slightly—the simplified model lacks a description of the sound spectral content and thus does not allow conversion to dB(A)—our comparison showed that the simplified description provides overestimated values, though never exceeding 3 dB. On the other hand, the computation time takes only seconds, which makes this description suitable for exploring a large set of solutions.

The model can be further refined if directivity information is available for the sound sources. Specifically, when the source is not omnidirectional, the radiated power is angle-dependent, and this dependency can be determined from the directivity pattern. The model can also take into account

- wind turbulence, which adds an omnidirectional frequency-dependent attenuation factor, and
- vertical wind and temperature gradients, which may increase the propagation effectiveness in some directions and introduce acoustic shadows in some others.[2]

The cumulative contribution of these atmospheric factors is difficult to predict for a generic time and location. However, some conditions usually observed together with high tides—for example, southeastern winds in autumn or winter—can be considered as a whole and distort the radiation pattern.

OPTIMAL PLACEMENT OF LOUDSPEAKERS

Auditory alert design guidelines commonly require that the acoustic stimulus must be about 15 dB above background noise to be clearly perceived.[3] Installation and maintenance costs make it impractical to install more than 8 to 12 loudspeakers in Venice's historic center and one each on separate islands such as Murano, Burano, and Lido. Although the average background noise in Venice is less than 60 dB, lower than in many other major cities, sound absorption makes it difficult to achieve the required loudness in dense areas with so few acoustic sources.

We therefore developed an optimization procedure to automatically determine better alert system configurations than the existing one.[4] A matrix represents a built-up area, and a pool P of matrix cells defines possible locations for the placement of acoustic sources. The noise maps that each source generates are assumed to be known, and they are computed using the simplified model. We therefore designed the optimization procedure to work independently of the propagation function. Because different functions can lead to different solutions, computation times also can vary.

Given n active sources ($n = 8$ in the current system), the problem consists of finding a subset of n locations in P that provides the best acoustic coverage of the whole area. This involves searching a broad tree of possible solutions defined by

$$\binom{|P|}{n} = \frac{|P|!}{(|P| - n)!\,n!}\ .$$

Operation Research has have developed a number of optimization frameworks suitable for our purpose including branch-and-bound techniques and simulated annealing or Monte Carlo strategies for traversing the search space. We used constraint logic programming over finite domains—in particular, the SICStus Prolog `clpfd` library (www.sics.se/sicstus). As the "Constraint Logic Programming" sidebar describes, CLP is a declarative programming paradigm that allows concise encoding of combinatorial minimization problems.

Constraint Logic Programming

Constraint logic programming[1,2] is a relatively new programming paradigm that is particularly well suited for encoding combinatorial minimization problems. CLP naturally merges two declarative paradigms: constraint solving and logic programming.

A *constraint* is, in general, a first-order formula that can be interpreted over various possible domains. For example, $1 \le X < X \wedge Y \wedge Y < 2$ is a constraint that is satisfiable over the domain **R** but not over **N**. CLP lets a programmer use different classes of constraints and domains to encode problems. For combinatorial problems, it is common to use finite domain constraints, namely arithmetic constraints between arithmetic expressions. The interpretation of variables, expressions, and constraints is over **Z**.

In classic logic programming,[3] a first nondeterministic phase generates a possible solution, then a second deterministic phase tests whether the solution is acceptable. If the search space grows exponentially with the input, this *generate-and-test* technique does not apply.

In contrast, CLP uses a *constrain-and-generate* technique in which a deterministic phase introduces a number of constraints, then a nondeterministic phase generates the solution space. This approach sensibly limits the number of potential solutions and makes it possible to exploit built-in algorithms, such as constraint propagation and branch-and-bound, and problem-solving heuristics.

For example, consider one possible encoding of an instance of the knapsack problem. A smuggler with a knapsack of size 49 units wishes to determine what combination of bottles of wine ($6 profit, 10 units), bottles of grappa ($10 profit, 17 units), and pieces of chocolate ($2 profit, 4 units) will maximize his profit.

The overall schema of a program of this form is

```
introduce_constraints(W,G,C,Profit),
labeling([maximize(Profit)],[W,G,C]).
```

where the predicate `introduce_constraints` is defined simply as

```
W #>= 0, G #>= 0, C #>= 0,
10*W+17*G+4*C #=< 49,
Profit #= 6 * W + 10 * G + 2 * C.
```

W stands for bottles of wine, *G* for bottles of grappa, and *C* for chocolate items. The built-in predicate *labeling* executes the constraint-driven search. The computed answer is $W = 1$, $G = 2$, $C = 1$, and $Profit = 28$.

References

1. J. Jaffar and M.J. Maher, "Constraint Logic Programming: A Survey," *J. Logic Programming,* vols. 19/20, 1994, pp. 503-581.
2. K. Marriott and P.J. Stuckey, *Programming with Constraints,* MIT Press, 1998.
3. L. Sterling and E. Shapiro, *The Art of Prolog,* 2nd ed., MIT Press, 1997.

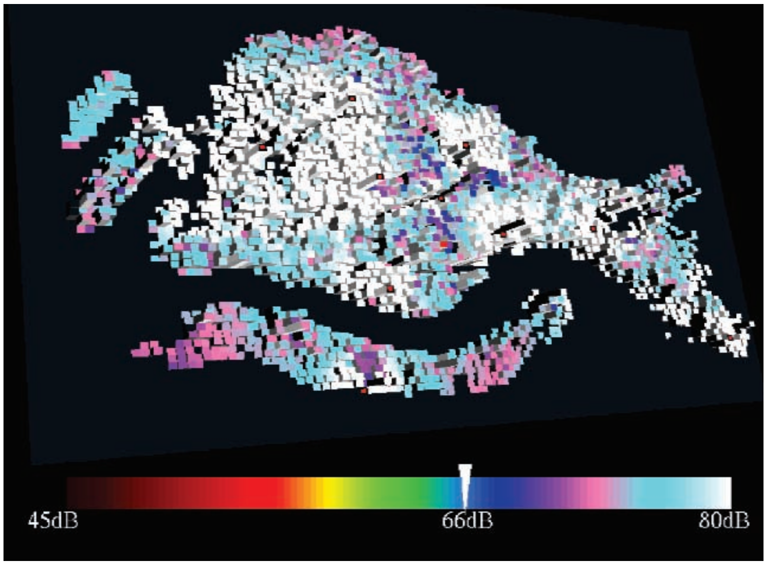


Figure 3. Proposed 10-point acoustic emission solution, according to maximum minimal intensity criterion. This optimization provides more uniform coverage than the current layout, especially in northwestern Venice, but it does not fulfill the recommended 75-dB threshold.

CLP includes a built-in constraint solver that uses branch-and-bound algorithms and other techniques to dramatically accelerate the search for solutions.

In our case, the main constraint was that the sound level in each grid cell had to exceed a given threshold. Another constraint required that only n sources could be active. A geometric constraint imposed a minimum Euclidean distance between two active sources.

We generated a pool P of 22 possible acoustic emission points ranging from 20 to 80 meters high, typically bell towers, somewhat uniformly distributed over the city area. Two points had to be included in all possible solutions because their locations were of primary importance. We then produced optimizations according to different criteria, such as minimum intensity variation or maximum minimal intensity, for 8 to 18 other emission points.

Figure 3 shows a 10-point solution optimized for maximum minimal intensity that provides more uniform acoustic coverage than the current layout, especially in northwestern Venice. The minimal measured sound intensity is 66 dB, and in most of the city it is well beyond 70 dB. Due to the constraint $n = 10$, the solution does not fulfill the recommended 75-dB threshold; this can only be satisfied by increasing n to 12 or more.

LARGE-SCALE ALERT DESIGN

Next, we designed the actual alert sounds. In contrast with a *warning*, which does not necessarily require action by the receiver, and an *alarm*, which requires immediate action, an *alert* requires action from the receiver at some point—that is, the auditory stimulus must be noticeable without being threatening.[3] Only a few studies, such as that by Benjamin Rubin,[5] have addressed the problem of auditory alert design in large-scale uncontrolled settings. Moreover, we faced numerous constraints not found in typical application areas such as an automotive environment or computer interfaces.

Sound types

Three types of sound commonly function as auditory alerts: speech, environmental (auditory icons), and abstract (earcons). Although speech signals would seem to offer the best way to communicate information about high tides, noise and echoes can inhibit speech intelligibility. In addition, because Venice attracts many foreign visitors, the signals should be recognizable to nearly everyone regardless of language or culture. Auditory icons are not suitable either, because many real environmental sounds are already present in city background noise. Abstract sounds are thus the most appropriate choice, especially if they resemble the siren sounds with which residents are familiar.

Information complexity

In addition to signaling a forthcoming high tide, the new auditory alerts needed to sonify the expected tide level to communicate the potential risk of flooding. To meet this requirement, we exploited the concept of *attensons,* attention-getting sounds often used in conjunction with verbal alerts.[3] According to this approach, the alert consists of two parts: an attenson, common to all alerts, that generally indicates a rising tide, and a signal that specifies the tide level. Studies on urgency mapping[3,6,7] provide various criteria that can be exploited for tide-level sonification.

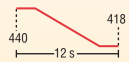
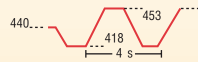
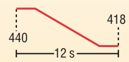
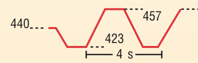
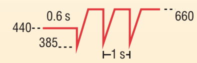
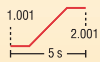
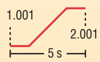
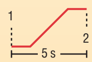
Physical constraints

The alert sounds must be audible at large distances—up to many hundreds of meters from the source. Environmental effects such as air absorption and multiple reflections can significantly alter the spectral content and time envelope of anechoic stimuli. Other relevant phenomena are the delay and spectral effects experienced when listening simultaneously to two or more sources located at different distances. Consequently, known design methods are not directly applicable to our case, necessitating additional care in choosing the parameters for controlling the urgency levels.

Training

Although the typical experimental setups used to evaluate alert signals and perceived urgency can draw upon direct training procedures, our target

Table 1. Control parameters for new alert sounds.

Parameter	Attenson + …	Tide level 1	Tide level 2	Tide level 3
Duration	5 seconds	12 seconds	14 seconds	16 seconds
Carrier frequency 1	440 Hz			
Carrier frequency 2	440 Hz			
Modulating frequency 1/ Carrier frequency 1		2.001	2.001	2.001
Modulating frequency 2/ Carrier frequency 2		2.001	2.001	2.001
Modulation index		2	2	2

population cannot. Instead, residents receive information mainly through local newspapers and ad hoc booklets. Consequently, some visual representation of the alert sounds is necessary for instructional purposes.

Retention

A disadvantage of abstract sounds is the difficulty associated with remembering large alert sets. Studies indicate that an individual can retain and recognize only four to seven sounds depending on the auditory dimensions.[3] In our study, retention is even more problematic because the alerts are so sporadic—typically 10 times per year—that residents must relearn the stimuli each time. Consequently, a small set of urgency levels is required, with dramatic differences between each level. The tide forecast models in current use can reliably predict tide levels within an error of ±10 cm, implying the need for at least three distinct tide-level alerts.

We designed our alert sounds using frequency modulation synthesis,[8] specifically with two modulated oscillators. This technique ensures the production of broadband spectra for appropriate parameter values, thereby minimizing masking problems. We initially selected fundamental frequencies in the 400-500 Hz range, which maximizes audibility at large distances.

Table 1 presents the control parameters for each alert sound, obtained as the sequence attenson + tide level n. The attenson's time evolution resembles that of the current siren sounds: increasing perceived pitch and opening of the spectrum. We produced the urgency levels through covariation of fundamental frequency, sound inharmonicity, and their temporal patterns.[7] These features are more robust to outdoor environmental effects than those characterizing Roy Patterson's pulse-burst approach,[6] such as pulse rate and interpulse interval.

We sonified level 1 with a slight pitch decrease. For level 2, we used a slow periodic pitch modulation (0.25 Hz) and slightly mistuned the two oscillators' carrier frequencies to add beatings and increase sound inharmonicity. We sonified level 3 with a faster (1 Hz) asymmetric pitch envelope, which ranges on a broader interval, and shifted the oscillators' envelopes by 0.3 seconds.

ALERT-DESIGN VALIDATION

To validate our designs, we conducted two types of tests: a rating test of the sounds and a matching test of the proposed visualizations of these sounds. The subjects did not receive formal training; rather, we informed them of the study's nature and gave them some simple instructions. Based on the results of these tests and other input, we refined our designs.

Rating test

The first test was conducted on 13 subjects, seven males and six females ranging from 19 to 51 years old. For the first test, we provided subjects with the following instructions:

> You will be presented with 12 alert sounds for high tide in Venice. Each lasts about 20 seconds, and is associated with a certain tide level. After each sound you are requested to write the tide level that matches that sound in your opinion. As a term of comparison, current alerts are used only for levels higher than 100 cm, while 140 cm is already an exceptionally high tide.

We played the three alert sounds once each in random order, and then played them again three times each in random order. The first three stimuli provided the subjects with a hidden training phase that we discarded from the results. Although we specified no strict range for the magnitude estimation task, the two anchors included in the instructions—100 and 140 cm—provided a strong

Table 2. Rating test results of 13 human subjects.

Urgency level	Mean judgment (cm)	Standard deviation (cm)
1	103.6	8.5
2	116.7	11.2
3	128.3	12.6

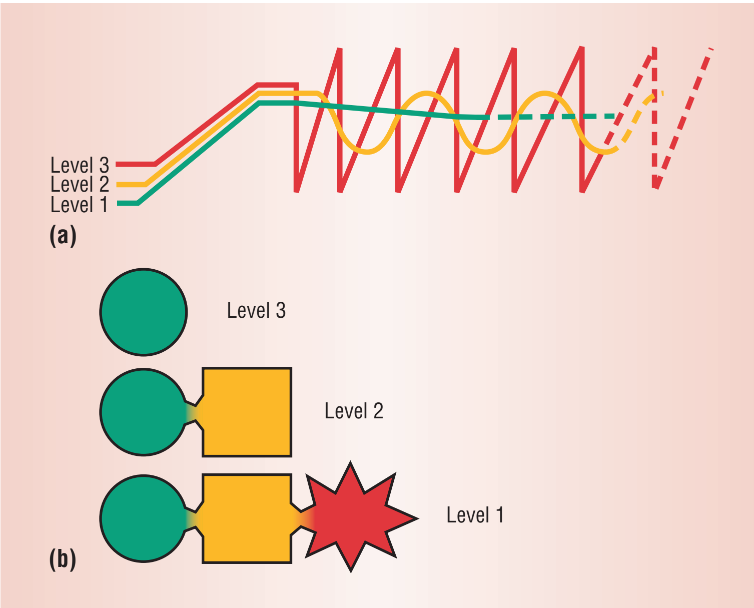


Figure 4. Visualizations of new alert sounds. (a) Some subjects confused the initial graphical design with waveforms. (b) A radically different visualization mapped each portion of the sound into geometric objects.

indication. Therefore, it was not necessary to introduce normalization to compensate for differences in the numerical scales used, and we could analyze the raw responses.

A repeated-measures analysis of variance on the mean magnitude estimates found the effect of the alert level to be statistically significant ($F = 43.67$, $p < 0.001$). Table 2 shows the mean magnitude estimates for each level and the corresponding standard deviation, which, incidentally, is roughly coincident with the uncertainty of forecasts. The ANOVA results showed no difference between Venetian and non-Venetian subjects.

Matching test

Figure 4a shows the results of our first effort to visualize the different sounds for local printed media. The three lines represent the pitch envelopes of each sound. Contour sharpness corresponds to inharmonicity. The initial pitch glides are juxtaposed to emphasize that the first part of the alert, the attenson, is common to all three sounds. The colors green, yellow, and red represent the urgency levels 1, 2, and 3, respectively.

To evaluate the effectiveness of this visualization strategy, we provided the 13 subjects with a color reproduction of Figure 4a and the following instructions:

> You will be presented with a 20-second alert sound for high tide in Venice. After listening to it, you are requested to write the tide level—low, medium, or high—that matches that sound in your opinion. This visual representation will help you in this task.

We conducted two iterations of the matching test, the second using a slightly modified graphical representation. We conducted two iterations of the matching test on 23 subjects, 14 males and nine females ranging from 26 to 49 years old. The second iteration used a slightly modified graphical representation. None of the subjects had previously participated in the rating test.

The test results and follow-up interviews revealed two main problems. First, subjects with musical or scientific training tended to interpret the lines as waveforms. Consequently, they associated the slightly inharmonic level 1 sound with the level 2 line on the sketch, which resembles a sinusoidal waveform, and the rougher level 2 and level 3 sounds with the sawtooth envelope. Second, when listening to a single stimulus, all the subjects tended to overestimate the urgency level of the first and second sounds.

Refining the design

These observations led us to redesign the auditory stimuli. We obtained the alert sound for a tide level *n* by juxtaposing the attenson with the sequence of level sonifications up to level *n*. We also created a radically different visualization that mapped each portion of the sound into geometrical objects with increasing "sharpness," as Figure 4b shows. This approach was inspired by Gestalt psychologists' observations on the intermodal character of the expressive qualities of communication.[9]

Field observations using an experimental emission station necessitated another acoustic redesign—namely, we had to shift the frequency content of stimuli downward and further differentiate the three alert timbres to overcome blending and filtering of signals as they propagate in open air.

Introducing various physical, economic, and social constraints can make a seemingly simple task, such as designing an acoustic alert system, quite complex. Tackling this problem on an urban scale required a broad perspective that encompassed logic programming, optimization, auditory display, geographical information systems, visualization, and user testing. Although our work addressed the precise requirements of the problem of high tides in Venice, most of the techniques and solutions we developed could be easily adapted to auditory alert systems for different hazards, such as toxic chemicals and severe weather, in other large-scale environments. ■

Acknowledgments

We thank Dario Bovo of Consorzio Venezia Ricerche, and Paolo Canestrelli and Leonardo Boato of the Centro Previsioni e Segnalazioni Maree, for their help and advice. Daniele Piccolo provided guidance with SoundPLAN acoustic simulations. Umberto Nicolao provided useful comments. We presented preliminary results of the work described in this article at the APPIA-GULP-PRODE 2002 Joint Conference on Declarative Programming and at the 2003 International Conference on Auditory Display ("*Acqua Alta a Venezia*: Design of a Urban Scale Auditory Warning System"; www.icad.org/websiteV2.0/Conferences/ICAD2003/paper/45%20Avanzini.pdf).

References

1. J.G. Roederer, *The Physics and Psychophysics of Music,* Springer-Verlag, 1995.
2. U. Ingard, "A Review of the Influence of Meteorological Conditions on Sound Propagation," *J. Acoustical Soc. of America,* vol. 25, no. 3, 1953, pp. 405-411.
3. N.A. Stanton and J. Edworthy, eds., *Human Factors in Auditory Warnings,* Ashgate, 1999.
4. F. Avanzini et al., "Optimal Placement of Acoustic Sources in a Built-Up Area Using CLP(FD)," *Proc. 2002 APPIA-GULP-PRODE Joint Conf. Declarative Programming*; www.dimi.uniud.it/dovier/PAPERS/agp02.pdf.
5. B.U. Rubin, "Audible Information Design in the New York City Subway System: A Case Study," *Proc. 1998 Int'l Conf. Auditory Display*; www.icad.org/websiteV2.0/Conferences/ICAD98/papers/RUBIN.PDF.
6. R.D. Patterson, "Auditory Warning Sounds in the Work Environment," *Philosophical Trans. Royal Soc. of London,* series B, vol. 327, no. 1241, 1990, pp. 485-492.
7. E.J. Hellier, J. Edworthy, and I. Dennis, "Improving Auditory Warning Design: Quantifying and Predicting the Effects of Different Warning Parameters on Perceived Urgency," *Human Factors,* vol. 35, no. 4, 1993, pp. 693-706.
8. C. Roads, *The Computer Music Tutorial,* MIT Press, 1996.
9. W. Köhler, *Gestalt Psychology,* Liveright, 1947.

Federico Avanzini *is a postdoctoral researcher in the Department of Information Engineering at the University of Padova. His research interests include sound/voice processing and human-computer interfaces, and audio rendering in multimodal and virtual reality systems. Avanzini received a PhD in computer science from the University of Padova. Contact him at avanzini@dei.unipd.it.*

Davide Rocchesso *is an associate professor in the Faculty of Mathematical, Physical, and Natural Science at the University of Verona. His research interests include vision, image, and sound processing. Rocchesso received a PhD in computer engineering from the University of Padova. He is a member of the IEEE Computer Society and the Acoustical Society of America. Contact him at davide.rocchesso@univr.it.*

Alberto Belussi *is an associate professor in the Faculty of Mathematical, Physical, and Natural Science at the University of Verona. His research interests include databases and information systems. Belussi received a PhD in computer engineering from the Polytechnic of Milan. Contact him at alberto.belussi@univr.it.*

Alessandro Dal Palù *is a PhD student in the Department of Mathematics and Computer Science at the University of Udine. His research interests include constraint programming, algorithms, and computational biology. Dal Palù received a Laurea in computer science from the University of Verona. He is a member of the Gruppo Ricercatori e Utenti Logic Programming (GULP). Contact him at dalpalu@dimi.uniud.it.*

Agostino Dovier *is an associate professor in the Department of Mathematics and Computer Science at the University of Udine. His research interests include constraint programming language design, programming with sets and multisets, computable set theory, Web-like databases and related graph algorithms, and computational biology. Dovier received a PhD in computer science from the University of Pisa. He is a member of GULP. Contact him at dovier@dimi.uniud.it.*