ACQUA ALTA A VENEZIA: DESIGN OF A URBAN SCALE AUDITORY WARNING SYSTEM

F. Avanzini†
Dip. Ing. dell'Informazione
University of Padova
Via Gradenigo 6/A
35131 Padova, Italy
avanzini@dei.unipd.it

D. Rocchesso
A. Belussi
Dip. Informatica
University of Verona
Strada le Grazie 15
37134 Verona, Italy
{*rocchesso;belussi*}*@sci.univr.it*

A. Dal Palù
A. Dovier
Dip. Matematica e Informatica
University of Udine
Via delle Scienze 206
33100 Udine, Italy
{*dovier;dalpalu*}*@dimi.uniud.it*

ABSTRACT

A new warning system for high tide in Venice has been designed to replace the existing network of electro-mechanical sirens. The project was divided into four sections: (i) optimal placement of loudspeakers via constraint logic programming, (ii) simulation and visualization of the acoustic field in the city, (iii) design of the warning sounds, (iv) validation of the warning sounds. This paper reports the strategies and results of all four project stages, with special emphasis on sound design and validation.

1. INTRODUCTION

The problem of high tide (*Acqua Alta*) in Venice is becoming more and more serious, due to recent changes in the morphology of the lagoon. High tide causes problems for pedestrians (rubber boots are always fashionable in Venice), public transportation (boats can not run under some bridges and alternative routes are activated), and shops (who try to defend themselves with fences). It is essential for the population to be informed in time, so that the appropriate measures can be taken[1]. A special office of the Municipality of Venice, the *Centro Previsioni e Segnalazioni Maree*, provides continuous tide forecast based on computational models, astronomical, and meteorological data. When a significantly high tide is expected, a network of electro-mechanical sirens is activated, usually anticipating the tide peak by a few hours.

A study for the realization of a new auditory warning system has been conducted in cooperation with the *Centro Previsioni e Segnalazioni Maree* and with the *Consorzio Venezia Ricerche*, with two main objectives: relocate the acoustic sources in order to ensure a more uniform coverage of the urban territory, and provide auditory information about the tide level (i.e., sonify different levels of gravity). The second point makes sense because the Municipality is considering the replacement of the electro-mechanical sirens with loudspeaker systems. If the environmental concerns raised by the loudspeakers will be overcome (a siren has less visual impact than a loudspeaker array), it will be possible to broadcast any kind of sound and replace the current threatening sounds, which are often associated with second world war air attacks. One may think that speech signals are well suited to inform about the tide level. However, since the beginning it was clear that the sounds should be audible from large distances in a variety of masking conditions, and that many different languages are spoken by the target population. Given these requirements, and considering that the inhabitants of Venice have learned to react to siren sounds over the years, we have decided to design non-speech signals somehow resemblant of the sounds used so far.

Figure 1: *The highest recorded tide in Venice (Nov. 4, 1966)*

The first part of the project was devoted to the optimal placement of the sound sources in the territory of Venice. Due to many physical, morphological and social constraints, this task is quite complex. The use of constraint logic programming for optimal placement was thoroughly described elsewhere [1], but it is briefly reviewed here. In this first phase we made extensive use of off-the-shelf acoustic simulation software (SoundPLAN), to validate the approximations introduced in the optimization procedure. A special visualization tool was written to allow comparative exploration of the solutions. The simulation and visualization activities are described in section 2. Section 3 discusses the optimal placement of sound sources. In section 4 we describe the design principles that were used to create the warning sounds, and the problems related to designing for sporadic and non-comparative listening. Section 5 reports about preliminary results of experiments designed to validate the proposed auditory warning system.

2. ACOUSTIC SIMULATIONS AND VISUALIZATION

The very first stage of the project was devoted to detailed analysis of the existing warning system. To this end, the Municipality of Venice provided digitalized geographical data (ArcView). A semi-automatic pre-processing step was developed to extract building

†Also with *CVR–Consorzio Venezia Ricerche.*

[1]Tourists seem to enjoy *Acqua Alta*, especially walking barefoot into it. But they want to be warned as well, in order to prepare films and cameras.

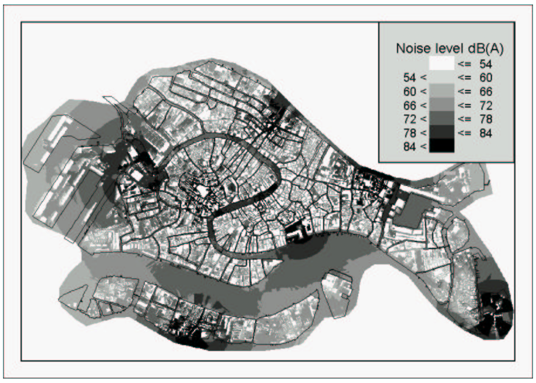


Figure 2: *The acoustic warning system currently in use.*

heights and areas with a reasonable degree of confidence. The data are structured in layers which include building polygons with associated heights, isle polygons, and water polygons (channels and lagoon). This information was passed to the software SoundPLAN [2] for sound pressure computation. The acoustic effects of map elements (reflection, refraction, absorption, shielding) were taken into account by the ray-tracing algorithm used in the simulations. The eight electro-mechanical sirens were modeled as omnidirectional point sources, and their spectral content was determined via FFT analysis of a steady-state portion of a siren sound.

Figure 2 depicts a noise map, obtained with SoundPLAN, of the current acoustic warning system. The coverage is far from uniform, and city areas with inadequate sound level are easily determined. The simulations used a $5\mathrm{m}$ grid step, a value which is larger than many Venetian channels and alleys. Even with such large step and with a reasonable number of ray reflections (four), the computation time of the ray-tracing algorithm was very large (in the order of days). Therefore, the use of general simulation tools such as SoundPLAN was found useful to produce a reliable image of the current situation and to assess the validity of the proposed solutions, but it is inappropriate both as an exploratory tool and as a routine to be embedded in optimization procedures.

A simplified acoustic description was then developed for this latter purpose. The main assumption is that the phases of signals coming from different sources are randomly mixed at the listening point (which is especially true in complex urban environments), so that the intensities of the component tones can be summed up constructively [3]. This assumption allows separate computations of the sound field of each source, regardless of the nature of the sound signals being emitted. In the simplified acoustic description, each source is modeled as a discrete variable function f which accounts for attenuation due to free-field propagation, additional air absorption, and shielding effect due to buildings.

Using the geographical data to compute the shielding effects, the function f was constructed as a simplified version of the international standard *VDI 2714/2720* [4]. This can be further refined by including directivity information for the sources, as well as wind effects. The actual form of the function f, however, does not affect the optimization procedure described in the next section.

3. OPTIMAL PLACEMENT OF THE LOUDSPEAKERS

In the literature of auditory warning design, a common requirement is that the acoustic stimulus must be about $15\mathrm{dB}$ above the background noise in order to be clearly perceived (see e.g. [5, chapter 4]). In the case here considered, for practical reasons no more than eight to ten sources can be used for the entire city[2]. Even though the average background noise in Venice is lower than in other cities (around $60\mathrm{dB}$), the loudness requirement is not trivially achieved with such a small number of acoustic sources, especially due to absorption in dense built-up areas.

In order to determine better system configurations than the existing one, an automatic procedure was designed. The procedure is explained in detail in [1], here only the basic ideas are reviewed. The built up area is represented by a matrix, and a pool P of locations (matrix cells) is defined, which collects the points where acoustic sources can be placed. The noise maps generated by each single source are assumed to be known, and they were computed using the simplified model presented in section 2. Given a number n of *active* sources (in the current system, $n = 8$), the problem consists in finding a subset of exactly n locations in the pool P that provides the best acoustic coverage of the whole area. This means that a broad ($\frac{|P|!}{(|P|-n)!n!}$) tree of solutions must be searched.

The problem was faced using Constraint Logic Programming over Finite Domains ($CLP(\mathcal{FD})$, see e.g. [6]). Using this approach the search for solutions is dramatically quickened by exploitation of the constraints involved in the problem. The main one consists in requiring the sound level in each grid cell to be greater than a given threshold. The second one states that only n sources must be active. A third geometric constraint imposes a minimum euclidean distance between two active sources. The procedure was implemented using the `clpfd` library of SICStus Prolog [7].

Possible emission points were chosen among high locations (typically, bell towers), ranging from $20\mathrm{m}$ to $80\mathrm{m}$ heights, somewhat uniformly distributed over the city area. As a result, a pool P of 22 points was constructed. Two points were forced to be included in all the possible solutions, since their locations are considered of primary importance. We produced optimizations according to different criteria, such as minimum intensity variation or maximum minimal intensity, for 8 to 10 emission points. Figure 3 shows a solution with 8 points, which is only slightly different from the current layout, but offers a much better coverage especially in the north-west side of the city. The minimal measured sound intensity is $62\mathrm{dB}$, and in most of the city it is well beyond $70\mathrm{dB}$. The visualization was obtained with an OpenGL based application, specifically written for this purpose.

4. DESIGN OF WARNINGS ON URBAN SCALE

This section reviews our approach to warning sound design. Few studies (see e.g. [8]) have addressed the problem of auditory warning design in large-scale non-controlled environments. In the situation examined here, additional constraints and problems are encountered, that are not found in typical application areas such as automotive environments or computer interfaces.

Types of sound. Three types of sound are most commonly used as auditory warnings: speech, abstract sounds (earcons), and environmental sounds (auditory icons). Speech signals are not suitable for our purposes. One reason is that they are more easily masked from background noise than non-speech broadband signals, another one is that the target population speaks many different languages. Environmental sounds are not suitable either, since

[2]We are referring here to the historical center. Separate islands (Murano, Burano, Lido, etc.) have one siren each.

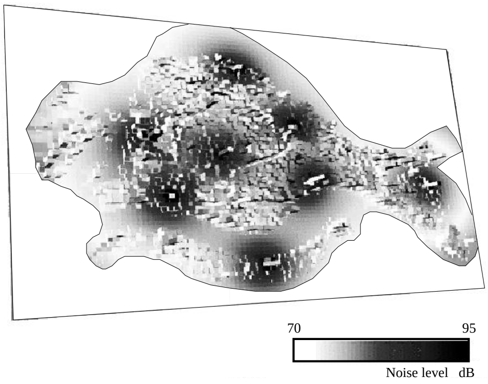


Figure 3: *Proposed solution with eight sound emission points.*

a variety of real environmental sounds are already present in the background noise. Abstract sounds are thus an appropriate choice.

Information complexity. The new auditory warnings must fulfill two main requirements: (i) signaling the approaching high tide, and (ii) sonifying the expected tide level. The current sounds fulfill only requirement (i). In order to include requirement (ii) in the new warnings, we have exploited the concept of *attensons* (attention-getting sounds), often used in conjunction with verbal warnings [5, chapter 14]. According to this idea, each sound signal is composed of two parts: the first one (the attenson) is common to all the warnings and fulfills requirement (i), while the second one conveys the information on tide level, i.e. requirement (ii). Studies on urgency mapping (see e.g. [9], [5, chapter 7], and [10]) have provided criteria that can be exploited in the level sonification.

Physical constraints. The warning sounds have to be audible within large distances (many hundreds of meters). Therefore environmental effects such as air absorption and multiple reflections can alter significantly spectral contents and time envelopes of the 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, design methods from the literature are not directly applicable to our case, and additional care must be taken in the choice of the parameters to be used for controlling the urgency levels.

Training. No direct training procedures for the target population can be implemented. Information is mainly provided through local newspapers and *ad hoc* booklets. This contrasts with any experimental set-up typically used in the evaluation of warning signals and perceived urgency. Consequently, some visual representation of the warning sounds must be developed, which can be used in order to instruct the population.

Retention. One disadvantage of abstract sounds is the difficulty associated with remembering large warning sets. Many authors have shown that, depending on the auditory dimensions, four to seven sounds can be retained and recognized by an individual [5, chapter 1]. In our case, the retention problem is even more relevant given that the warning sounds are heard sporadically (typically less than ten times in year). This means that the stimuli are hardly memorized, and are rather re-learned at each new high tide alarm. Consequently, a small set of urgency levels must be used, with dramatic differences between each level. The current tide forecast models can reliably predict tide levels within an error of $\pm 10\mathrm{cm}$, in practice this implies that warnings for three distinct tide levels are needed.

The warning sounds have been designed using FM (frequency modulation) synthesis, specifically with two modulated oscillators. This technique ensures that broadband spectra are produced for appropriate parameter values, thus minimizing masking problems. The fundamental frequencies are chosen in the 400–$500\mathrm{Hz}$ range, which maximizes audibility at large distances.

The parameter values for each sound are schematically given in table 1. The warning sound for a tide level n is obtained as (Attenson + Level n). Note in particular that the time evolution of the attenson is designed to resemble that of the current siren sounds (increasing perceived pitch and opening of the spectrum). Urgency levels were produced through covariation of a few sound features. The pulse-burst approach described by Patterson [9] was not taken, since parameters such as pulse rate and interpulse interval are not robust to outdoor environment effects. We selected the fundamental frequency, the sound inharmonicity, and their temporal patterns as the more relevant features for the sonification of tide levels: these are more robust to environmental effects and are known to be relevant to perceived urgency [10].

Level 1 was sonified with a slight pitch decrease; a slow periodic pitch modulation ($0.25\mathrm{Hz}$) was used in level 2, moreover the carrier frequencies of the two oscillators were slightly mistuned in order to add beatings and increase the sound inharmonicity; level 3 was obtained with a faster ($1\mathrm{Hz}$) and asymmetric pitch envelope, which ranges on a broader interval. Note also that in level 3 the envelopes for the two oscillators are shifted in time ($0.3\mathrm{s}$).

5. VALIDATION AND DISCUSSION

Rating. A first listening test was conducted with 13 subjects, seven males and six females aged 19–51 years, both Venetians and non-Venetians. No training was given to the subjects, they were told the nature of the study and given the instructions: *"You will be presented with 12 warning sounds for high tide in Venice. Each lasts about 20s, and is associated with a certain tide level (in cm). After each sound you are requested to write the tide level (in cm) that matches that sound in your opinion. As a term of comparison, current warnings are used only for levels higher than 100cm, while 140cm is already an exceptionally high tide"*.

After the first sound, each stimulus was presented to the subject when he/she had finished writing. The twelve stimuli were presented with the following internal organization (not known to the subject): first the three warning sounds were played once each in random order, then they were played again three times each in random order. This way, the first three stimuli provided the subject with a "hidden" training phase. Table 2 shows that the subjects ranked the stimuli in the correct order. No differences were found between venetian and non-venetian subjects.

Matching. As already discussed, some form of visualization of the acoustic warnings must be found in order to train the target population through local newspapers and *ad hoc* booklets. The

Level	Mean Judgment (cm)	St. deviation (cm)
1	103.6	8.5
2	116.7	11.2
3	128.3	12.6

Table 2: *Rating test: mean levels and standard deviations.*

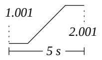
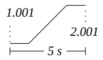
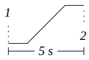
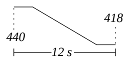
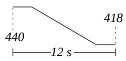
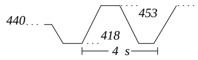
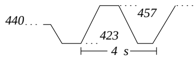
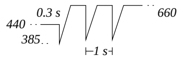
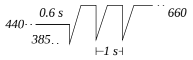
	D (s)	C1 (Hz)	C2 (Hz)	M1/C1	M2/C2	I
Attenson	5	440	440			
Level 1	12			2.001	2.001	2
Level 2	14			2.001	2.001	2
Level 3	16			2.001	2.001	2

Table 1: *Control parameters for the warning sounds: D stands for duration (in s), C1 and C2 for the first and second carrier frequencies (in Hz), respectively, M1 and M2 for the first and second modulating frequencies (in Hz), respectively, and I for the modulation index.*

picture in figure 4(a) was designed as a first attempt. The three polylines represent the pitch envelopes of each sound. Increasing inharmonicity is somewhat rendered through increasing contour sharpness. The initial pitch glides are juxtaposed in order to emphasize that the first part common to the three sounds. Urgency level is rendered using different colors for the envelopes, namely green, yellow, and red, for level 1, 2, and 3 respectively.

A second listening test was being set up in order to evaluate the effectiveness of this visualization strategy. Again, no training was provided to the subjects, they were given a color reproduction of figure 4(a) and the following instructions: *"You will be presented with a 20s warning sound for high tide in Venice. After listening to it, you are requested to write the tide level (low, medium, high) that matches that sound in your opinion. This visual representation will help you in this task."*.

Discussion. Two iterations of the matching test have been already conducted, the second one used a sligthly modified graphical representation. Results are not satisfactory, and interviews have revealed two main problems. *(1)* Subjects with musical/scientific training tend to interpret the graphical sketches in terms of waveforms and consequently associate level 1 (sound) to level 2 (sketch) because it is only slightly inharmonic (and thus resemblant of a sinusoidal waveform), while rougher sounds are associated with the sawtooth envelope. *(2)* When listening to a single stimulus, all the subjects tend to overestimate the urgency level of the first and second sounds.

Following these observations, we are considering the redesign the auditory stimuli using an "incremental approach: the warning sound for a tide level n is obtained as (Attenson + $\sum_{i=1}^{n}$ Level n).

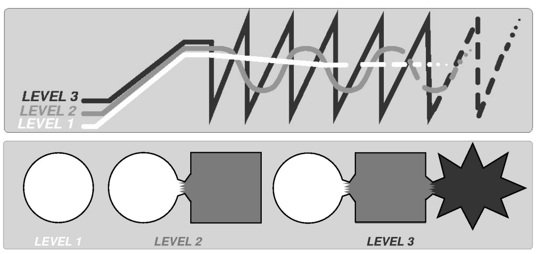


Figure 4: *Visualization of the warning sounds; (a) first approach (level 1 green, level 2 yellow, level 3 red); (b) second approach (circle green, square yellow, star red).*

At the same time a totally different visualization approach is being experimented, where each portion of the sound is mapped into geometrical objects with increasing sharpness (see figure 4(b)).

Acknowledgments

We thank Dario Bovo, director of *CVR–Centro Venezia Ricerche*, and Paolo Canestrelli, director of *Centro Previsioni e Segnalazioni Maree*, for their help. Daniele Piccolo (*CVR*) provided guidance to the use of SoundPLAN and contributed to the acoustic simulations.

6. REFERENCES

[1] F. Avanzini, A. Belussi, A. Dal Palù, A. Dovier, and D. Rocchesso, "Optimal Placement of Acoustic Sources in a Built-up Area using $CLP(\mathcal{FD})$," in *Proc. APPIA-GULP-PRODE 2002, Joint Conf. on Declarative Programming*, Madrid, 2002, pp. 139–154.

[2] Braunstein + Berndt GmbH SoundPLAN LLC, "Soundplan home page," `http://www.soundplan.com/`.

[3] J. G. Roederer, *The Physics and Psychophysics of Music*, Springer Verlag, New York, 1995.

[4] R. Tonin, "Modelling and predicting environmental noise," Tech. Rep., R. Tonin & Associates Pty Ltd, `http://www.rtagroup.com.au`.

[5] N.A. Stanton and J. Edworthy (eds.), *Human factors in auditory warnings*, Ashgate Publishers: Aldershot, 1999.

[6] K. Marriott and P. J. Stuckey, *Programming with Constraints*, MIT Press, 1998.

[7] Swedish Institute for Computer Science, "Sicstus prolog home page," `http://www.sics.se/sicstus/`.

[8] B. U. Rubin, "Audible information design in the new york city subway system: A case study," in *Proc. Int. Conf. Auditory Display (ICAD'98)*, Glasgow, Nov. 1998.

[9] R. D. Patterson, "Auditory warning sounds in the work environment," *Phil. Trans. Royal Society of London*, vol. B 327, pp. 485–492, 1990.

[10] 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, pp. 693–706, 1993.