Eco-sistema

by Max Jacob & Enrico Di Ienno

Overview

Eco-sistema is a computer driven sound installation thought for public urban spaces, both indoor and outdoor, having an approximately circular or n-sided polygonal shape ranging from 10x10m to 30x30m.

The aim of the *eco-sistema* is to be a barometer and regulator of the acoustical ambiance it is installed in, interacting in a possibly constructive way with the sound produced by people and machines, but, at the same time, trying to call attention on the responsibility we all have in taking care of our daily life sound-scape.

The installation consists of six loudspeakers and three microphones connected to a computer. The computer analyses the incoming data (from the microphones) and generates audio controlled by the outcomes of this analysis.

There are two parallel mechanisms running in *eco-sistema*: the **crickets** and the **harmonic series**. The first is a virtual emulation of six crickets reacting to the acoustical ambiance, the chirping of each one is emitted by one of the loudspeakers. The second is a set of three harmonic series (each one emitted by a speaker pair) tuned on the ambiance *fundamental frequency*, and whose purpose is to smooth down the more or less abrupt changes in the site's acoustical ambiance by enveloping each sonic event in a kind of harmonic cloud.

The installation is designed in a way not to compromise any acoustical communication occurring in the venue, and can thus be set up in super-markets, railway stations etc. without affecting their regular operation.

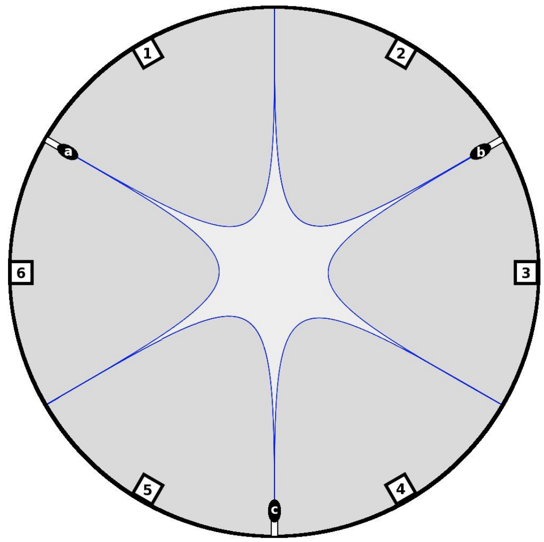
Eco-sistema runs on Linux and has been realized using exclusively free third party or self-written software.

We see *eco-sistema* as an example of a new generation of *ambient music*, not being composed once and propagated afterwards by some loudspeakers, but rather composing itself while it is playing, constantly adapting to and inspired by the environment it has to interact with.

Description

General setup

The ideal venue for *eco-sistema* is a circular space, where speakers and microphones can be positioned as in the following plan:



The speakers are labelled with numbers (1-6), and the microphones by letters (*a*, *b*, *c*). The darker grey areas surrounded by a blue line in front of each speaker are the *sectors* of that speaker.

Speakers and microphones are installed ideally at a height of about half the space radius, but it can be less than this.

All speakers and microphones are connected to a computer.

The shape of the venue can be quite different than the ideal circular form: it can be a square or any not too uneven polygon where width and depth don't differ too much.

The audio captured by the microphones is analysed by the *eco-sistema* software in order to detect significant events and general tendencies needed for controlling the generated audio. This one consists of two parallel mechanisms: the **crickets** and the **harmonic series**, described in next paragraphs.

The crickets

The tendential chirping frequency

Each of the six speakers emits the chirping of a virtual cricket. The chirping can be fast (about two chirps per second) or slow (up to one chirp each three seconds), depending on the daytime. The curve which determines the mapping between daytime and (tendential) chirping frequency should be redesigned in accordance to the venue's daily life cycle.

The chirping rhythm

Any sonic event occurring in the sector of a speaker/cricket causes a hesitation (longer pause) in the chirping of that cricket, proportionally to the intensity of the event. The crickets get used to the acoustical ambiance they are in, and sonic events are always considered within their context: in a loud ambiance a sonic event must be very loud to be considered, while in a silent ambiance a little noise can cause several seconds of hesitation.

The overall result is that there are six chirping crickets, each with its own *personality* depending on what happens around it. Since the chirping rhythm is controlled by *real life* data captured from the ambiance, it never sounds mechanical or repetitive.

Stops and restarts

When a very loud event occurs, that is captured by all microphones (what means that it has an acoustical impact on the whole venue), the crickets suddenly stop, all together and for a long while (at least 2-3 minutes). The threshold above which such an event is considered loud enough is computed adaptively depending on the acoustical ambiance of the venue, avoiding crickets to be almost always silent or never stop chirping.

In order to restart, each cricket analyzes its own sector, and restarts only if during the past 3 minutes there has been a continuous reduction of sonic activity. This will happen, sooner or later, for all crickets. The effect is that, after a pause, the crickets restart timidly one by one beginning where noise has calmed down.

The output level

Cricket chirping covers mainly frequencies above 4.5kHz, what means that it does not mask significantly human voices which carry information mainly on a lower spectrum. However we want the crickets to be audible, but on the background. The output level is thus kept quite low, but is adapted in real time to the acoustical ambiance and raised slightly as the ambiance becomes louder, in order to ensure audibility.

The harmonic series

A harmonic series is a set of sinusoids whose frequency is a whole multiple of a base frequency called the *fundamental*. Any (tuned) musical instrument playing a note produces such a harmonic series having the note as fundamental. The timbre of different instruments depends (mainly) on the relative level of the sinusoids composing the series.

Tuning

The system tries to detect continuously the *fundamental frequency* in the acoustical ambiance, and uses this as the base frequency for the three harmonic series it generates. In most cases this will be almost a random value, since most ambiances are mainly a-periodic (noisy). However if it happens that a loud periodic signal occurs (as the whistle of a train, for example), it will be recognized and used to tune the harmonic series.

Behaviour

Each harmonic series is emitted by a pair of adjacent loudspeakers, and considers the sonic activity within the two sectors of those speakers. The behaviour is based on a running mean of the sound spectrum within the pertinent sectors during the past 30 seconds. If in the instantaneous spectrum (the current ambiance) there is something missing compared to this mean on some frequency band, each harmonic series fills up the ambiance by raising the level of the harmonics falling within that band.

The result is a sound cloud which changes constantly seeping into the *holes* left free by the acoustical ambiance while retiring when the holes are filled up by the ambiance itself. Another aspect is that when a sonic event occurs causing the harmonic series to retire, this will raise the reference mean as well. Thus, as the event stops, the harmonic series will detect a *larger hole* and react with more intensity.

Propagation

As has been said, each series is emitted by a speaker pair, and all six speakers emit, although with different levels, the same harmonics. This makes that the position of the source is much less intelligible in comparison to what happens with the crickets. The three series are not perceived so much as three distinct phenomena, but rather as a single sound cloud whose spatial distribution changes over time, more or less rapidly depending on what happens in the venue.

Moreover, what the listener hears exactly depends from his precise spatial position, because sinusoids emitted from different speakers can be either in phase (raised) or in phase opposition (attenuated) in the room of a few centimetres. As the listener moves around, these phase effects cause rapid and unpredictable shimmering in the perceived sound.

The feedback

Eco-sistema has a quite efficient anti-feedback system, what means that it is able to subtract from the captured audio the sound coming from the loudspeakers (to which we don't want it to react).

However this system is not (and can not) be perfect. Thus, if the venue is very very silent, *eco-sistema* ends up in *listening* a little bit to itself. This minimal interaction between the machine and itself causes the installation to be *alive* even in the extreme case of a perfectly silent venue.

Purpose

In Italian *eco* means both «echo» and the prefix «eco-». *Eco-sistema* could thus be translated in English to *echo-system*, summarizing the fact that it is a system reacting to the sound produced in the venue (like an echo does) and that it does its best to constantly find a balance with its environment (as organisms in an ecosystem do).

Eco-sistema reveals by its reactions phenomena and aspects that usually keep unnoticed, as medium term changes in the sonic activity, the spectral distribution in space, the occurring of sudden sonic events etc.. But at the same time it tries to cooperate constructively with the (usually not well balanced) sound of the venue, trying to produce a less stressing and annoying overall result. This is exactly the opposite (cooperation vs. conflict) to what happens usually when music (propagated by loudspeakers or by the earphone of an mp3 player) is used to overcome an existing sonic ambiance judged as unpleasant.

What makes the sound of the installation interesting from a strictly perceptive point of view, is that each sound is controlled by data captured from the real life, reflecting its fluctuations, hidden structures, regularities and oddities. The visitor can just pass by without paying much attention to the installation, or he can stop and listen to it as one listens to a water stream, to singing birds or the whistling of the wind.

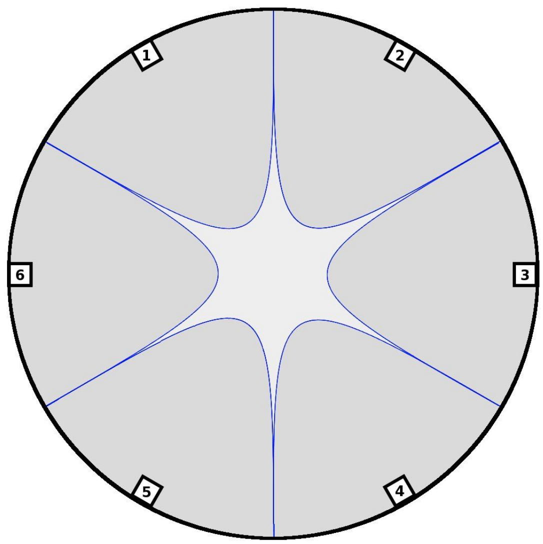
Eco-sistema has been designed, according to its “ecological” philosophy, to blend smoothly in any venue, without disturbing anybody and without compromising the intelligibility of any acoustical communication signal. This is the main reason why we have chosen to emulate crickets, which operate on quite high frequencies outside of the domain of normal human communication. The harmonic series, on their side, retire in front of any sound higher than the (adaptively computed) norm, causing them to be silenced, for example, by a communication propagated through a loudspeaker or by somebody shouting a message to someone else at the other side of a street. In general the output level is kept low enough that people can talk each other without problems, at least as far as the venue isn't too noisy anyway.

We hope that *eco-sistema* can be the first example of a new generation of *ambient music*, not being composed once and propagated afterwards by some loudspeakers, but rather composing itself while it is playing, constantly adapting to and inspired by the environment it has to interact with.

Technical details

Position of speakers and microphones

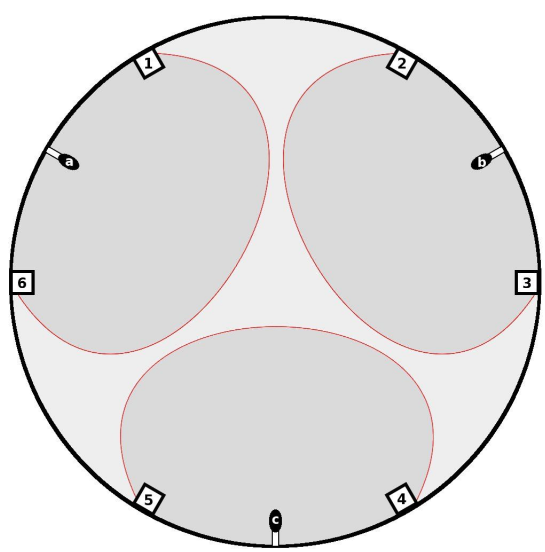
As has been mentioned, the ideal location for *eco-sistema* is a circular space, where speakers can be positioned as in the following plan:



Each speaker is here represented by a white box with black border, labelled with a number. The principle is that each speaker covers a sixth of the location area. In the diagram the sector of each speaker is represented by a grey area surrounded by a blue line. Each speaker should be the most audible one for a visitor stationing inside its sector. This is achievable by installing the speakers at a reasonable height (ideally about half the location radius, but it can be less than this) and by carefully registering their vertical angle in order to adhere as much as possible to the foreseen scheme.

The shape of the location can be quite different than the ideal circular form: it can be a square or any not too uneven polygon, as long as it can be cut into six approximately triangular sectors of more or less the same area.

Following plan shows the position of the microphones (labeled with “a”, “b”, and “c”):



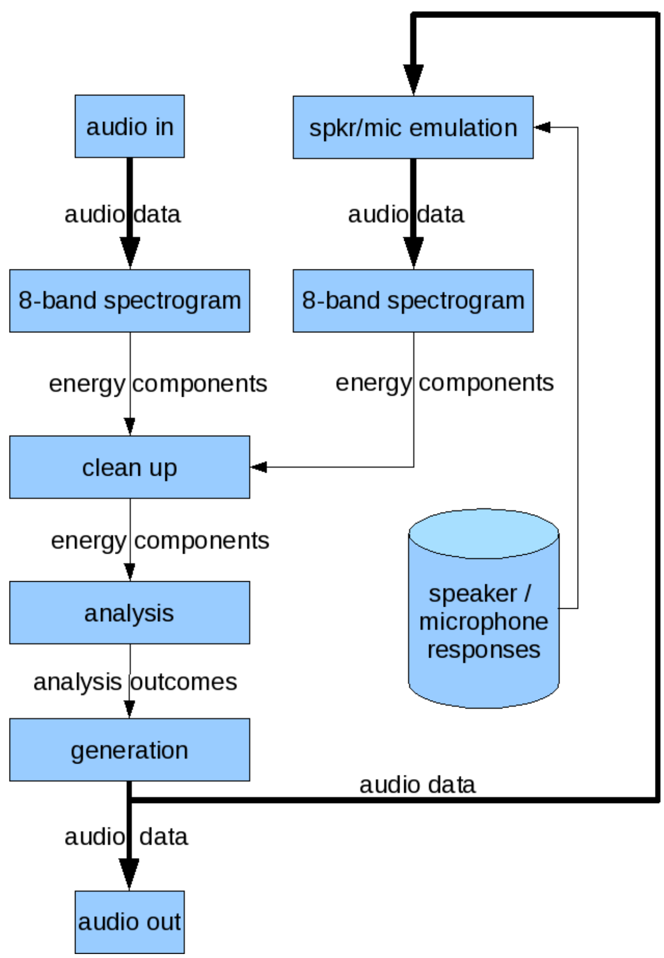
Each microphone covers roughly one third of the location area. Each sound produced within the sector of a microphone should be captured by that microphone at a higher level than by the others. The microphones should be installed at the same height than the speakers, and the angle must be carefully registered in order to get as close as possible to the above scheme.

The analysis of the captured audio data

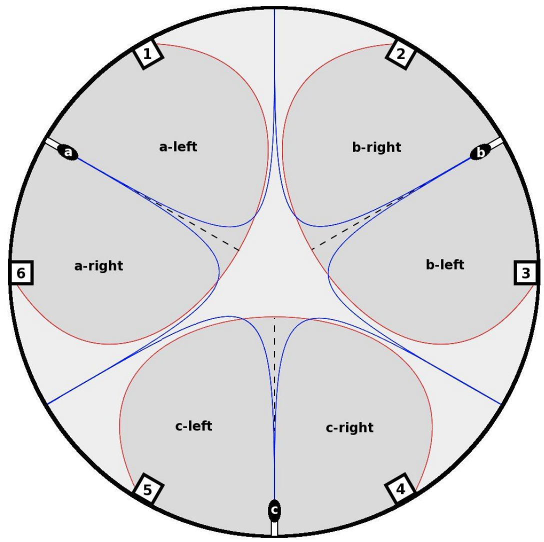
First of all there is a little paradox: e*co-sistema* reacts acoustically to the acoustical ambiance it is installed in. But we do not want *eco-sistema* to react to itself, what would be the case in a simplistic implementation, since the acoustical ambiance contains as well *eco-sistema*'s contribution. To handle this, we have written a quite efficient *anti-feedback system* that is able to subtract the computer generated sound from the one captured by the microphones.

This is achieved by a piece of software that performs, at set-up time, an analysis on how a sound propagated by each single loudspeaker is captured by each single microphone in that specific venue and set-up. This information is used afterwards, at run time, to predict continuously *eco-sistema*'s contribution to the captured signals, making it possible to clean them up.

This task, which is commonly very hard to perform in real life situations (outside a research lab), is made quite easy (and thus well working) in *eco-sistema* because what is analysed is not the sampled audio signal, but the energy components of this signal on eight frequency bands (a little broader than one octave each) at a much lower rate (about 40 Hz). Following diagram summarizes this logic:



Once the signal has been cleaned up, a first analysis stage generates six energy component signals obtained by comparing the three incoming ones. This step allows to *locate* sonic events as coming from either the left or right area of each microphone sector, corresponding roughly to the speaker sectors (in blue in the diagram):



The accuracy of the analysis heavily depends on a careful set-up and tuning before the installation is started. The advantages in using three microphones instead of six are basically: less computation need (instead of simulating 36 speaker/mic combinations, we only have to simulate 18), less overall feed-back (3 microphones pick up less loudspeaker sound than 6) and, obviously, less microphones to rent, position and connect. Furthermore the six channel information is needed only by the crickets, while the harmonic series operate on speaker pairs and are fed directly with the data coming from one microphone (1 and 6 from *a*, 2-3 from *b*, 4-5 from *c*).

Subsequent analysis stages compute (often time-weighted) statistics on the 8 frequency bands of each signal independently. This allows to detect events better than simply analysing the overall energy, since distinct events will probably cause only some bands to raise, which is less noticeable (and might be masked by another sound going away) just looking at the overall energy.